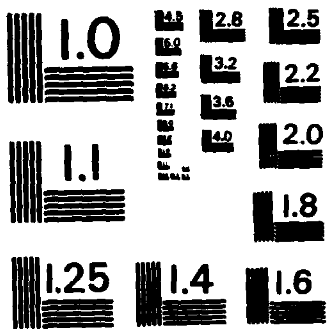
AD-A160 322 REMARKS ON THE FOUNDATIONS OF MEASURES OF DEPENDENCE (U) NORTH CAROLINA UNIV AT CHAPEL HILL CENTER FOR STOCHASTIC PROCESSES R C BRADLEY ET AL. JUN 85 TR-105 1/1

UNCLASSIFIED AFOSR-TR-85-0877 F49620-82-C-0009 F/G 12/1 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

REMARKS ON THE FOUNDATIONS OF MEASURES OF DEPENDENCE

by
Richard C. Bradley
Włodzimierz Bryc
Svante Janson

TECHNICAL REPORT 105

June 1985

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED

1b. RESTRICTIVE MARKINGS:

2a. SECURITY CLASSIFICATION AUTHORITY:

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE:

3. DISTRIBUTION/AVAILABILITY OF REPORT: unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S): Technical Report 105

5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 85-0877

6a. NAME OF PERFORMING ORGANIZATION: Center for Stochastic Processes

6b. OFFICE SYMBOL (If applicable):

7a. NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research

6c. ADDRESS (City, State and ZIP Code):
University of North Carolina
Statistics Dept., 321 PH 039A
Chapel Hill, NC 27514

7b. ADDRESS (City, State and ZIP Code):
Bolling Air Force Base
Washington, DC 20332

8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR

8b. OFFICE SYMBOL (If applicable):

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F49620 82 C 0009

8c. ADDRESS (City, State and ZIP Code):
Bolling AFB
Washington, DC 20332

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2304	A5	

11. TITLE (Include Security Classification): REMARKS ON THE FOUNDATIONS OF MEASURES OF DEPENDENCE

12. PERSONAL AUTHOR(S): Richard C. Bradley, Włodzimierz Bryc, Svante Janson

13a. TYPE OF REPORT: Technical

13b. TIME COVERED: FROM 9/84 TO 8/85

14. DATE OF REPORT (Yr., Mo., Day): June 1985

15. PAGE COUNT: 52

16. SUPPLEMENTARY NOTATION:

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Measures of dependence, norms, covariance.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This paper is a study of several aspects of measures of dependence, the various comparisons between them, and their foundation as norms of the bilinear form covariance.

Additional keywords: stochastic processes; triangles, inequalities, Banach space; random variables, vector analysis; Hilbert space.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL: Brian W. Woodruff, Maj, USAF

22b. TELEPHONE NUMBER (Include Area Code): 767-5027

22c. OFFICE SYMBOL: NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REMARKS ON THE FOUNDATIONS OF MEASURES OF DEPENDENCE

by

Richard C. Bradley
Department of Mathematics
Indiana University
Bloomington, Indiana 47405
USA

and

Center for Stochastic Processes
Department of Statistics
University of North Carolina
Phillips Hall 039A
Chapel Hill, North Carolina 27514
USA

Włodzimierz Bryc
Instytut Matematyki
Politechnika Warszawska
00-661 Warszawa
POLAND

Svante Janson
Department of Mathematics
Uppsala University
Thunbergsvägen 3
752 38 Uppsala
SWEDEN

The work of Richard C. Bradley was partially supported by NSF Grant No. DMS 84-01021 and AFOSR Grant No. F49620 82 C 0009.

0. INTRODUCTION

Some notations: Before we start, let us quickly list some of the notations and conventions that will be used.

When an expression of the form a_b is a subscript or exponent, it will sometimes be written as a(b) for typographical convenience.

The characteristic function (indicator function) of an event A in a probability space will be denoted by I_A or $I(A)$.

The complement of an event A in a given probability space will be denoted by A^c.

The fraction 0/0 is always interpreted to be 0.

For any number p, $1 \le p \le \infty$, the conjugate exponent will be denoted by p' $(1 \le p' \le \infty)$; that is, $1/p + 1/p' = 1$.

For a given bounded signed measure μ, the total variation of μ (on the whole measurable space) will be denoted by $\mathrm{var}(\mu)$.

The zero element of a given Banach space will be denoted simply by 0.

For $n = 1,2,3,\ldots$,

$$[1,\infty]^n := \{(p_1,\ldots, p_n): 1 \le p_k \le \infty \text{ for all } k = 1,\ldots, n\}.$$

Let $(\Omega, \mathcal{M}, P)$ be a fixed probability space. By a "measure of dependence" we mean any function d mapping pairs of sub-σ-fields of $\mathcal{M}$ into $\overline{\mathbb{R}}_+ := \mathbb{R}_+ \cup \{0\} \cup \{\infty\}$ and satisfying the following two natural requirements:

(0.1) For any two σ-fields $\mathcal{F}$ and $\mathcal{G} \subset \mathcal{M}$, $d(\mathcal{F},\mathcal{G}) = \sup d(\mathcal{F}_0,\mathcal{G}_0)$ where the sup is taken over all pairs of finite σ-fields $\mathcal{F}_0 \subset \mathcal{F}$ and $\mathcal{G}_0 \subset \mathcal{G}$; and

(0.2) $d(\mathcal{F},\mathcal{G}) = 0$ if and only if $\mathcal{F}$ and $\mathcal{G}$ are independent σ-fields.

As a consequence of eqn. (0.1), d is "increasing": If $\mathcal{F}_1 \subset \mathcal{F}$ and $\mathcal{G}_1 \subset \mathcal{G}$ then $d(\mathcal{F}_1,\mathcal{G}_1) \le d(\mathcal{F},\mathcal{G})$.

This paper is mainly concerned with "dominations" of measures of dependence in the following sense:

We say that a measure of dependence d_1 is "dominated" by another measure of dependence d_2 (and we write $d_1 \prec d_2$) if there exists a function $\Phi: \overline{\mathbb{R}}_+ \to \overline{\mathbb{R}}_+$ such that $\Phi(0) = 0$, Φ is continuous at 0, and $d_1 \leq \Phi(d_2)$ (i.e. $d_1(F,G) \leq \Phi(d_2(F,G))$ for all pairs of σ-fields F and $G \subset M$).

If $d_1 \prec d_2$ and $d_2 \prec d_1$, then we say that d_1 and d_2 are "equivalent".

Note that if (F_n, G_n), $n = 1,2,\ldots$ is a sequence of pairs of σ-fields such that $d_2(F_n,G_n) \to 0$ as $n \to \infty$, then for any $d_1 \prec d_2$ we have that $d_1(F_n,G_n) \to 0$. Conversely, if $d_1(F_n,G_n) \to 0$ as $n \to \infty$ for every sequence of pairs of σ-fields such that $d_2(F_n,G_n) \to 0$, then $d_1 \prec d_2$. (To see why, consider the function $\Phi(t) := \sup\{d_1(F,G) : d_2(F,G) \leq t\}$.)

Stronger types of "domination" may be defined by imposing conditions on Φ. For example, Φ linear implies the strong condition $d_1 \leq Cd_2$. In all cases of domination studied in [3] and in this paper, $\Phi(t)$ may be taken as a power t^δ, with $\delta > 0$, for small t. This implies that if $d_2(F_n,G_n) \to 0$ exponentially fast then so does $d_1(F_n,G_n)$. This type of property is of interest in connection with certain mixing conditions, but we shall not pursue this further in the present paper.

Here and throughout the paper, for measures of dependence d_1 and d_2, the equation $d_1 = d_2$ means that $d_1(F,G) = d_2(F,G)$ for all pairs of σ-fields F and G, and the equation $d_1 \leq d_2$ means that $d_1(F,G) \leq d_2(F,G)$ for all F,G. Thus $d_1 \leq d_2$ is technically a stronger statement than $d_1 \prec d_2$.

Here are some examples of known measures of dependence that we shall be interested in later on. Prior to each one, the name of the corresponding mixing condition for stochastic processes is given:

Strong mixing,

(0.3) $\alpha(F,G) := \sup|P(A\cap B) - P(A)P(B)|, \quad A\in F,\ B\in G$

ϕ-mixing,

(0.4) $\phi(F,G) := \sup|P(B|A) - P(B)|, \quad A\in F,\ B\in G,\ P(A)>0$

ψ-mixing,

(0.5) $\psi(F,G) := \sup\left|\frac{P(A\cap B)}{P(A)P(B)} - 1\right|, \quad A\in F,\ B\in G,\ P(A)P(B)>0$

ρ-mixing,

(0.6) $\rho(F,G) := \sup|\mathrm{Corr}(X,Y)|, \quad X\in L_2(F),\ Y\in L_2(G),$

X real, Y real

Absolute regularity ("weak Bernoulli"),

(0.7) $\beta(F,G) := \sup \tfrac{1}{2}\sum_{i=1}^{I}\sum_{j=1}^{J}|P(A_i\cap B_j) - P(A_i)P(B_j)|$

where this sup is taken over all pairs of partitions $\{A_1,\ldots,A_I\}$ and $\{B_1,\ldots,B_J\}$ of Ω such that $A_i\in F$ for all i and $B_j\in G$ for all j. To put this another way, $\beta(F,G) = (\tfrac{1}{2})\cdot\mathrm{var}(P_{F\times G} - P_F\times P_G)$ where $P_{F\times G}$ is the restriction to $F\times G$ of the measure on $\Omega\times\Omega$ induced by P and the diagonal mapping $\omega\mapsto(\omega,\omega)$ and P_F, P_G are respectively the restrictions to F and G of the measure P.

ρ-mixing again (see [2], [3, Theorem 1.1(ii)], or [4]),

(0.8) $\lambda(F,G) := \sup\frac{|P(A\cap B) - P(A)P(B)|}{[P(A)P(B)]^{1/2}}, \quad A\in F,\ B\in G,$

$P(A)P(B)>0$

Some simple dominations between the above measures of dependence are given by simple inequalities:

$\alpha(F,G) \le \beta(F,G) \le \phi(F,G) \le \psi(F,G)$

$\alpha(F,G) \le \lambda(F,G) \le \rho(F,G)$

The domination

$\rho(F,G) \le 2\cdot\phi^{1/2}(F,G)$

is well known (see e.g. [9, p. 309, Theorem 17.2.3]), and it has been improved independently by Denker and Keller [6, p. 506, line 2 and p. 516, line -8]: $\rho(F,G) \le 2\cdot\max\{\phi(F,G), \phi(G,F)\}$, and by Peligrad [13, p. 462, eqn. (4)]: $\rho(F,G) \le 2\cdot\phi^{\frac{1}{2}}(F,G)\cdot\phi^{\frac{1}{2}}(G,F)$. Another domination, taken from [3, Theorem 1.1], is as follows:

$$\rho(F,G) \le C\cdot\lambda(F,G)\cdot[1 - \log\lambda(F,G)]$$

where C is a universal constant.

The present paper continues and complements [3]. Now [3] was motivated by Rosenblatt's [15, Chapter 7] use of the Riesz convexity (interpolation) theorem to compare mixing conditions on Markov chains. Also, Lifshits [11, Lemma 1] used the Riesz-Thorin interpolation theorem in order to establish a moment inequality involving several measures of dependence. The idea of [3] was to use operator theory to try to develop a unified approach to the study of a large class of measures of dependence, and to prove "general" domination results for this class. The main point in [3] was to consider a particular bilinear form:

$$B(f,g) := E(fg) - (Ef)(Eg) \tag{0.9}$$

defined for simple functions f and g such that f is F-measurable and g is G-measurable (where F and G are, say, given σ-fields $\subset M$). Various norms of the bilinear form B define a large class of measures of dependence. Following [3], for $1 \le p,q \le \infty$, for given σ-fields F and G, we shall be interested in the (p,q)-norm of B on $F \times G$, namely

$$R_{p,q}(F,G) := \sup \frac{|E(fg) - (Ef)(Eg)|}{\|f\|_p \|g\|_q} \tag{0.10}$$

f simple, complex-valued, F-measurable

g simple, complex-valued, G-measurable

Later on, we shall also be interested in a version of (0.10) using r.v.'s f and g taking their values in Banach spaces. We shall also consider for fixed $0 \le r,s \le 1$,

$$(0.11) \qquad \alpha_{r,s}(\mathcal{F},\mathcal{G}) := \sup \frac{|P(A\cap B) - P(A)P(B)|}{[P(A)]^r[P(B)]^s}, \quad A \in \mathcal{F}, \quad B \in \mathcal{G}$$

which is a variant of $R_{1/r,1/s}(\mathcal{F},\mathcal{G})$ in which only indicator functions are being used. In the terminology of [16], $\alpha_{r,s}$ is simply the "restricted" $(1/r,1/s)$-norm of the bilinear form B.

Of the six measures of dependence in eqns. (0.3)-(0.8) above, five can immediately be fit into this framework, as follows:

$$\alpha = \alpha_{0,0}, \quad \phi = \alpha_{1,0}, \quad \psi = \alpha_{1,1}, \quad \lambda = \alpha_{\frac{1}{2},\frac{1}{2}}, \quad \rho = R_{2,2}$$

(The last equality $\rho = R_{2,2}$ follows from [18, p. 512, Theorem 1.1] and a trivial calculation; restriction to real, mean-zero r.v.'s would have no effect on $R_{p,q}$ in the special case $p = q = 2$). The remaining one, namely β in eqn. (0.7), will be expressed in terms of (0.9) using B-valued (Banach space valued) r.v.'s in Section 2.2 later on.

When we wish to consider a version of (0.10) for just real-valued r.v.'s, we shall use the notation

$$(0.12) \qquad R_{p,q}(\mathcal{F},\mathcal{G};\mathbb{R}) := \sup \frac{|E(fg) - (Ef)(Eg)|}{\|f\|_p\|g\|_q}$$

$$f \text{ simple, real-valued, } \mathcal{F}\text{-measurable}$$

$$g \text{ simple, real-valued, } \mathcal{G}\text{-measurable}$$

A trivial calculation yields

$$(0.13) \qquad R_{p,q}(\mathcal{F},\mathcal{G};\mathbb{R}) \le R_{p,q}(\mathcal{F},\mathcal{G}) \le 4\cdot R_{p,q}(\mathcal{F},\mathcal{G};\mathbb{R})$$

(and one can lower the 4 somewhat). When we don't wish to mention $\mathcal{F}$ and $\mathcal{G}$

explicitly, the measure of dependence $R_{p,q}(\cdot,\cdot;\mathbb{R})$ will be written simply as $R_{p,q}(\mathbb{R})$.

In eqns. (0.10) and (0.12) the restriction to simple functions will be convenient, but it is obviously a stronger restriction than necessary.

The following proposition is well known and quite elementary:

Proposition 0.1. If $1 \le p,q \le \infty$, and F and G are σ-fields, then the following two statements hold:

(i) $$R_{p,q}(F,G) = \sup \frac{\|E(f|G) - Ef\|_{q'}}{\|f\|_p}$$

f simple, complex-valued, F-measurable

(ii) $$R_{p,q}(F,G;\mathbb{R}) = \sup \frac{\|E(f|G) - Ef\|_{q'}}{\|f\|_p}$$

f simple, real-valued, F-measurable

Using this proposition, some results in Rosenblatt [15, Chapter 7] can be transcribed into our language, including the following statement from [15, p. 211, Theorem 1]: The measures of dependence $R_{p,p'}$ (and $R_{p,p'}(\mathbb{R})$), $1 < p < \infty$, are equivalent.

Remark 0.2. There is a nice connection between Lorentz spaces and the measures of dependence $\alpha_{r,s}$. In this remark we shall summarize some key points of this connection; the reader is referred to e.g. [1] [8] [19] for the details on Lorentz spaces. In fact $\alpha_{r,s}$ can be considered as a Lorentz space norm of the bilinear form B in (0.9). Recall (see e.g. [1]) that the Lorentz space L_{p1}, $1 \le p < \infty$, may be defined as the completion of the set of simple functions with the norm $\|f\|_{p,1} := \inf \sum_i |a_i| \cdot P(A_i)^{1/p}: \sum_i a_i I_{A(i)} = f\}$. This definition makes sense also for $p = \infty$, and $L_{\infty,1}$ equals L_∞ with an "equivalent norm" (see [1]). If $1 \le p < \infty$, then the dual space of L_{p1} equals $L_{p',\infty} := \{f: \sup_{t>0} t^{1/p'} P(|f| > t) < \infty\}$. Let $p = 1/r$ and $q = 1/s$. Then $\alpha_{r,s}$ equals the norm of B on $L_{p1} \times L_{q1}$ or, equivalently, if $q < \infty$, the norm of the linear operator $E(\cdot|G) - E(\cdot)$ from $L_{p1}(F)$ into $L_{q',\infty}(G)$. It is obvious that, with $r = 1/p$ and $s = 1/q$, one has $\alpha_{r,s}(F,G) \le R_{p,q}(F,G)$. On the other hand, $L_{p1} \supset L_{p_1}$ for every $p_1 > p$, and thus $\alpha_{r,s}(F,G) \ge C \cdot R_{p_1,q_1}(F,G)$ if $p_1 > 1/r$ and $q_1 > 1/s$. (Here the constant C depends on r,s,p_1,q_1.)

Remark 0.3. All of the measures of dependence d that are studied in this paper have, in addition to eqns. (0.1) and (0.2), three other nice properties.

(a) $d(\bigvee_{n=1}^{\infty} F_n, \bigvee_{n=1}^{\infty} G_n) = \lim_{N\to\infty} d(\bigvee_{n=1}^{N} F_n, \bigvee_{n=1}^{N} G_n)$.

(b) The measure d is given by a well defined formula that is applicable on any probability space.

(c) For finite σ-fields F and G the quantity $d(F,G)$ does not depend on the particular choice of F and G or on the underlying probability space (Ω,M,P), but depends only on the matrix $(P(A_i \cap B_j): 1 \le i \le I, 1 \le j \le J)$ where $A_1,\ldots, A_I$ (resp. $B_1,\ldots, B_J$) are (in any order) the atoms of F (resp. G). Thus (by eqn. (0.1)) measures of dependence can be formulated in terms of properties of doubly stochastic matrices. Some of the "domination" results obtained in this paper may be closely related to matrix inequalities (of which there is already a huge theory).

Remark 0.4. Keeping in mind Remark 0.3(b) for the measures of dependence that we study, the statement "d_1 fails to dominate d_2" will be used to mean that there exists a probability space on which d_1 fails to dominate d_2. (It is easily seen that in such cases d_1 will fail to dominate d_2 on the probability space [0,1] with Lebesgue measure, or on any other atomless probability space.) To prove that d_1 fails to dominate d_2 it suffices to show that for different choices of pairs of σ-fields F and G, on different choices of probability spaces, one can have $d_1(F,G)$ arbitrarily small without $d_2(F,G)$ being (arbitrarily) small (for the same F and G). For then by taking the product of countably many such probability spaces, one can obtain a single probability space on which d_1 fails to dominate d_2 (i.e. on which there are σ-fields F_n and G_n, $n = 1,2,3,\ldots$ such that $d_1(F_n, G_n) \to 0$ as $n \to \infty$ but $d_2(F_n, G_n)$ fails to converge to 0).

In [3] a detailed analysis of dominations between $\alpha_{r,s}$ and $R_{p,q}$ was given in the "lower triangle", $r + s \leq 1$ resp. $1/p + 1/q \leq 1$. In particular it was shown that among these measures of dependence in the lower triangle there are only four equivalence classes. Section 1 of this paper will complement that analysis. There, in Section 1.1 it will be shown that [3, Theorem 3.6] (one of the main tools for the study of measures of dependence in [3]) is within a multiplicative constant of being sharp. In Section 1.2 it will be shown that for $\alpha_{r,s}$ and $R_{p,q}$ in the "upper triangle", $r + s > 1$ resp. $1/p + 1/q > 1$, there are no equivalencies except for the well known one $\alpha_{1,1} = R_{1,1}$. In Section 1.3 a short proof of [3, Theorem 2.2] is given, using a three-step "reiteration" technique. Section 1.4 contains sharp domination results for the case where one of the two σ-fields is finite with only two atoms, each having probability ½.

Section 2 is devoted to measures of dependence based on B-valued (Banach space valued) r.v.'s. In Section 2.1 a preliminary discussion is given. In Section 2.2 the dependence coefficient β (in eqn. (0.7)) will be fit into this B-valued scheme, and it will also be shown that if $p > 1$ and $q > 1$ then $R_{p,q}$ fails to dominate β. In Section 2.3 we drop an unnecessary assumption made in [3, Theorem 4.2].

SECTION 1: THE SCALAR CASE

In Section 1 here we study measures of dependence based on random variables taking scalar (i.e. real or complex) values, namely the measures of dependence $R_{p,q}(F,G)$ and $R_{p,q}(F,G;\mathbb{R})$ from eqns. (0.10) and (0.12). Because of eqn. (0.13) we really only need to discuss one of these two (families of) measures; following [3] we shall discuss $R_{p,q}(F,G)$ (which uses complex r.v.'s). Of course the measures $\alpha_{r,s}$ will also be part of this discussion, which is intended to complement [3].

SECTION 1.1: SHARPNESS OF A DOMINATION RESULT

Section 1.1 here is the only part of this paper in which measures of dependence between more than two σ-fields will be included. In [3], Theorem 3.6 was a key tool for comparing some measures of dependence between two or more σ-fields. Here, for any given choice of parameters meeting the specifications in [3, Theorem 3.6], we shall show that [3, Theorem 3.6] is within a constant factor of being sharp. The sharpness (in the same sense) of [3, Theorem 2.1] follows as an indirect consequence. (Otherwise a sharper version of Theorem 2.1 would lead to a sharper version of Theorem 3.6 by the same proof, and this would contradict the example given below.) The nature of the example given below is such that it also confirms the sharpness (in the same sense) of [3, Theorem 4.1(vi)].

Before stating the result, let us recall some terminology from [3]. Suppose $(\Omega,\mathcal{M},P)$ is a probability space, and $F_1, F_2,\ldots, F_n$ are σ-fields $\subset \mathcal{M}$. Suppose $B: S(F_1) \times\ldots\times S(F_n) \to \mathbb{C}$ is an n-linear form, where $S(F)$ denotes the set of (equivalence classes of) F-measurable complex-valued simple functions. Suppose $\underline{p} := (p_1,\ldots, p_n)$ where $1 \le p_k \le \infty$ for all $k = 1,\ldots, n$. Then we define (as in [3]),

$$(1.1.1) \qquad \|B\|_{\underline{p}} := \sup \frac{|B(f_1, \ldots, f_n)|}{\|f_1\|_{p_1} \cdot \ldots \cdot \|f_n\|_{p_n}}$$

$$f_k \in S(F_k) \text{ for all } k = 1, \ldots, n$$

$$(1.1.2) \qquad d_{\underline{p}}(B) := \sup \frac{|B(I_{A_1}, \ldots, I_{A_n})|}{\|I_{A_1}\|_{p_1} \cdot \ldots \cdot \|I_{A_n}\|_{p_n}}$$

$$A_k \in F_k \text{ for all } k = 1, \ldots, n$$

<u>Theorem 1.1.1</u>. Suppose $n \geq 2$ is an integer and $1 \leq p_k \leq \infty$, $k = 1, \ldots, n$. Define $\underline{p} := (p_1, \ldots, p_n)$. Then there exists a constant C depending only on $\underline{p}$ such that the following holds:

For each t, $0 < t < 1$, there exists a probability space (Ω, M, P) and σ-fields $F_1, \ldots, F_n \subset M$, and an n-linear form $B: L_\infty(F_1) \times \ldots \times L_\infty(F_n) \to \mathbb{C}$ (namely $B(f_1, \ldots, f_n) := E(f_1 \cdot \ldots \cdot f_n) - \Pi_{k=1}^n Ef_k)$ such that $d_{\underline{p}}(B) \leq t$ and $\|B\|_{\underline{p}} \geq Ct(1 - \log t)^c$ where $c := \Sigma_{k \in K} 1/p_k'$ with $K := \{k: 1 < p_k < \infty\}$.

The constant exponent c here also depends only on $\underline{p}$ and is exactly the same as in [3, Theorem 3.6]. (Thus [3, Theorem 3.6] is within a constant factor of being sharp. See also the Addendum on p. 52 of this report.)

<u>Proof</u>. If $p_k \in \{1, \infty\}$ for all $k = 1, \ldots, n$ then the theorem becomes trivial. So we assume that $1 < p_k < \infty$ for at least one k.

For each p, $1 < p < \infty$, and each v, $0 < v \leq \frac{1}{2}$, define the function $G_{v,p}: [0,1] \to [0,1]$ as follows:

$G_{v,p}(x) := \min\{x, vx^{1/p}\}$ for $0 \leq x \leq \frac{1}{2}$

$G_{v,p}(x) := G_{v,p}(1 - x)$ for $\frac{1}{2} \leq x \leq 1$

Note that G is concave and increasing on $[0,\frac{1}{2}]$ and hence G is concave on $[0,1]$.

For $1 < p < \infty$ and $0 < v \le \frac{1}{2}$ define the function $g_{v,p}\colon [0,1] \to [-1,1]$ by $g_{v,p}(x) := \frac{d}{dx} G_{v,p}(x)$. That is,

$$g_{v,p}(x) := \begin{cases} 1 \text{ if } 0 < x < v^{p'} \\ v\cdot(1/p)\cdot x^{-1/p'} \text{ if } v^{p'} < x < \frac{1}{2} \\ -v\cdot(1/p)\cdot(1-x)^{-1/p'} \text{ if } \frac{1}{2} < x < 1 - v^{p'} \\ -1 \text{ if } 1 - v^{p'} < x < 1 \end{cases} \tag{1.1.3}$$

($g_{v,p}$ is not defined at $x = 0$, $v^{p'}$, $\frac{1}{2}$, $1 - v^{p'}$, 1.) Then $g_{v,p}$ is non-increasing, $\int_0^1 g_{v,p}(x)\,dx = 0$, and $|g_{v,p}(x)| \le 1$ for all x at which $g_{v,p}(x)$ is defined.

If $0 < v \le \frac{1}{2}$ and $p = 1$ or ∞, define the function $g_{v,p}\colon [0,1] \to [-1,1]$ by

$$g_{v,p}(x) := \begin{cases} 1 \text{ if } 0 < x < \frac{1}{2} \\ -1 \text{ if } \frac{1}{2} < x < 1 \end{cases} \tag{1.1.4}$$

In this case too, $g_{v,p}$ is non-increasing, $\int_0^1 g_{v,p}(x)\,dx = 0$, and $|g_{v,p}(x)| \le 1$.

The following integral will be used later on. If $0 < v \le \frac{1}{2}$ and $1 < p < \infty$, then

$$\int_0^{\frac{1}{2}} |g_{v,p}(x)|^{p'} dx = v^{p'}[1 + (1/p)^{p'}(\log \tfrac{1}{2} - p' \log v)] \tag{1.1.5}$$

Now let us get to the main part of the argument.

Henceforth m denotes Lebesgue measure on $[0,1]$.

Let t, $0 < t < (\frac{1}{2})^n$ be arbitrary but fixed. (There is no loss of generality in specifying $0 < t < (\frac{1}{2})^n$ rather than $0 < t < 1$ as in the statement of Theorem 1.1.1.)

Define v, $0 < v < \frac{1}{2}$, by

$$v^{\operatorname{card}\{k:\ 1 < p(k) < \infty\}} = t \tag{1.1.6}$$

Define the probability space (Ω,M,P) as follows:
$\Omega: = [0,1]^n: = [0,1] \times [0,1] \times\ldots\times [0,1]$. M is the σ-field of Borel subsets of Ω. P is defined by

$$P(A_1 \times\ldots\times A_n) = \prod_{k=1}^{n} m(A_k) + \prod_{k=1}^{n} \int_{A_k} g_{v,p_k}(x)dx$$

for all Borel subsets $A_1,\ldots, A_n \subset [0,1]$. (Recall the inequality $|g_{v,p_k}(x)| \le 1$ mentioned above.)

Note that since $\int_0^1 g_{v,p}(x) = 0$, each of the marginal distributions of p is uniform on $[0,1]$.

For each $k = 1,\ldots, n$ let F_k be the σ-field generated by the k^{th} coordinate in Ω. Define the n-linear form $B: L_\infty(F_1) \times\ldots\times L_\infty(F_n) \to \mathbb{C}$ as follows:

$$B(f_1,\ldots, f_n): = E(f_1 \cdot\ldots\cdot f_n) - \prod_{k=1}^{n} Ef_k$$

<u>Proof that</u> $d_{\underline{p}}(B) \le t$:

Suppose $D_k \in F_k$, $k = 1,\ldots, n$ with $P(D_k) \ne 0$. For each k, represent D_k by $D_k: = [0,1] \times\ldots\times [0,1] \times B_k \times [0,1] \times\ldots\times [0,1]$ (where the k^{th} coordinate-set B_k is a Borel subset of $[0,1]$). Then

$$|B(I_{D_1},\ldots, I_{D_n})| = \prod_{k=1}^{n} |\int_{B_k} g_{v,p_k}(x)dx|$$

$$\le \prod_{k=1}^{n} [\int_0^{m(B_k)} g_{v,p_k}(x)dx]$$

$$= [\prod_{\{k:1 < p_k < \infty\}} G_{v,p_k}(m(B_k))] \cdot [\prod_{\{k\,:\,p_k=1 \text{ or } \infty\}} [\int_0^{m(B_k)} g_{v,p_k}(x)dx]]$$

$$\le [\prod_{\{k:1 < p_k < \infty\}} v\cdot(m(B_k))^{1/p_k}] \cdot [\prod_{\{k:p_k = 1\}} m(B_k)] \cdot [\prod_{\{k:p_k = \infty\}} 1]$$

$$= v^{\operatorname{card}\{k:1<p(k)<\infty\}} \cdot \prod_{k=1}^{n} \|I_{D_k}\|_{p_k}$$

$$= t \cdot \prod_{k=1}^{n} \|I_{D_k}\|_{p_k}$$

by eqns. (1.1.3), (1.1.4), and (1.1.6) and the fact that for each $0 < v < \frac{1}{2}$, $1 \le p \le \infty$ the function $g_{v,p}$ is non-increasing (as noted earlier) and odd-symmetric. Thus $d_{\underline{p}}(B) \le t$.

<u>Proof that</u> $\|B\|_{\underline{p}} \ge Ct(1 - \log t)^c$ (for some constant C depending only on $\underline{p}$). (Here c is as in the statement of Theorem 1.1.1.)

Define the r.v.'s $f_1, \ldots, f_n$ as follows:

$$f_k(x_1, \ldots, x_n) := g_{v,p_k}(x_k) \quad \text{if} \quad p_k = 1 \text{ or } \infty$$

$$f_k(x_1, \ldots, x_n) := [\operatorname{sign} g_{v,p_k}(x_k)] \cdot |g_{v,p_k}(x_k)|^{p_k'/p_k} \quad \text{if} \quad 1 < p_k < \infty$$

To shorten the notation below, we write $f_k(x)$ instead of $f_k(x_1, \ldots, x_{k-1}, x, x_{k+1}, \ldots, x_n)$. Note that for each k, $f_k \in L_\infty(F_k)$.

Now

$$B(f_1, \ldots, f_n) = \prod_{k=1}^{n} \int_0^1 f_k(x) g_{v,p_k}(x)\,dx$$

For each k such that $p_k = 1$ or ∞, $\int_0^1 f_k(x) g_{v,p_k}(x)\,dx = 1$ and $\|f_k\|_1 = \|f_k\|_\infty = 1$.

For each k such that $1 < p_k < \infty$, $\int_0^1 f_k(x) g_{v,p_k}(x)\,dx = 2 \cdot \int_0^{\frac{1}{2}} |g_{v,p_k}(x)|^{p_k'}\,dx$ and $\|f_k\|_{p_k} = 2^{1/p_k} \cdot [\int_0^{\frac{1}{2}} |g_{v,p_k}(x)|^{p_k'}\,dx]^{1/p_k}$.

Hence

$$B(f_1, \ldots, f_n) = \Big[\prod_{\{k:1<p_k<\infty\}} 2^{1-1/p_k} [\int_0^{\frac{1}{2}} |g_{v,p_k}(x)|^{p_k'}\,dx]^{1-1/p_k}\Big] \cdot \prod_{k=1}^{n} \|f_k\|_{p_k}.$$

For each k such that $1 < p_k < \infty$,

$$\int_0^{1/2} |g_{v,p_k}(x)|^{p_k'}\, dx \geq C_k \cdot v^{p_k'} (1 - \log t)$$

by eqn. (1.1.5), where $C_k := \min\{1 + (1/p_k)^{p_k'} \log \tfrac{1}{2},\ (1/n)\cdot(1/p_k)^{p_k'}\}$. (Note that C_k is positive and depends only on $\underline{p}$.)

Hence by eqn. (1.1.6), $B(f_1,\ldots, f_n) \geq C\cdot t\cdot(1 - \log t)^c \cdot \prod_{k=1}^{n} \|f_k\|_{p_k}$ where c is as in Theorem 1.1.1 and $C := \prod_{\{k:1<p_k<\infty\}} C_k^{1/p_k'}$ (which is positive and depends only on $\underline{p}$).

Now trivially the supremum of $|B(f_1,\ldots, f_n)|/(\|f_1\|_{p_1}\cdot\ldots\cdot\|f_n\|_{p_n})$ is the same over $(f_1,\ldots, f_n) \in S(F_1)\times\ldots\times S(F_n)$ as over $L_\infty(F_1)\times\ldots\times L_\infty(F_n)$. Hence $\|B\|_{\underline{p}} \geq C\cdot t\cdot(1 - \log t)^c$. This completes the proof.

Remark 1.1.2. [3, Theorem 3.6] holds for $n = 1$ as well as for $n \geq 2$, by a simple argument similar to the proof of [3, Theorem 2.1 or 2.2]. Also, [3, Theorem 3.6] is within a constant of being sharp for $n = 1$, as one can show with a slight modification of the above example.

SECTION 1.2: DOMINATIONS IN THE UPPER TRIANGLE

The following definition will be needed: For any point (r,s) such that $0 < r,s \leq 1$, and $r + s > 1$, define $Q(r,s)$ to be the closed (convex) quadrilateral region with vertices (0,0), (1,0), (r,s), and (0,1).

Remark 1.2.1. Statements (a)-(g) below give a complete picture of what dominations (or equivalencies) exist for any two of the various measures of dependence $\alpha_{r,s}$ and $R_{p,q}$ with $0 \leq r,s \leq 1$ and $1 \leq p,q \leq \infty$. This whole picture can be pieced together from [3, Theorem 4.1 and Remark 4.1] and Propositions 1.2.2-1.2.6 given below; to do so, one should keep in mind the trivial facts that for any two σ-fields F and G, $\alpha_{r,s}(F,G) \leq R_{1/r,1/s}(F,G)$

and $\alpha_{r,s}(F,G)$ and $R_{1/r,1/s}(F,G)$ are each non-decreasing as r or s increases. (In particular, the problem mentioned in the last sentence of [3, Remark 4.1] is solved.) The restrictions $0 \le r,s \le 1$ and $1 \le p,q \le \infty$ are to be implicitly understood.

(a) The measures of dependence $\alpha_{r,s}$, $r+s<1$, and $R_{p,q}$, $1/p+1/q<1$, are equivalent; and they do not dominate any of the other measures of dependence $\alpha_{r,s}$ and $R_{p,q}$.

(b) The measures of dependence $\alpha_{r,1-r}$, $0<r<1$, and $R_{p,p'}$, $1<p<\infty$, are equivalent; and they dominate $\alpha_{r,s}$, $r+s<1$, and $R_{p,q}$, $1/p+1/q<1$; but they do not dominate any of the other measures of dependence $\alpha_{r,s}$ and $R_{p,q}$.

(c) $\alpha_{1,0}$ and $R_{1,\infty}$ are equivalent. $\alpha_{0,1}$ and $R_{\infty,1}$ are equivalent. Neither $\alpha_{1,0}$ dominates $\alpha_{0,1}$, nor vice versa. $\alpha_{1,0}$ and $\alpha_{0,1}$ each dominate the measures of dependence $\alpha_{r,s}$, $r+s<1$, $R_{p,q}$, $1/p+1/q<1$, $\alpha_{r,1-r}$, $0<r<1$, and $R_{p,p'}$, $1<p<\infty$. Neither $\alpha_{1,0}$ nor $\alpha_{0,1}$ dominates any of the measures of dependence $\alpha_{r,s}$, $r+s>1$, or $R_{p,q}$, $1/p+1/q>1$.

(d) $\alpha_{1,1} = R_{1,1}$; and these measures of dependence dominate all of the other ones $\alpha_{r,s}$ and $R_{p,q}$.

(e) If $0<r_0,s_0<1$ and $r_0+s_0>1$, then α_{r_0,s_0} dominates $\alpha_{r,s}$ and $R_{1/r,1/s}$ for all $(r,s) \in Q(r_0,s_0)$, except that α_{r_0,s_0} fails to dominate $\alpha_{1,0}$, $R_{1,\infty}$, $\alpha_{0,1}$, $R_{\infty,1}$, or $R_{1/r_0,1/s_0}$. For such (r_0,s_0), R_{r_0,s_0} dominates $\alpha_{r,s}$ and $R_{1/r,1/s}$ for all $(r,s) \in Q(r_0,s_0)$, except that R_{r_0,s_0} fails to dominate $\alpha_{1,0}$, $R_{1,\infty}$, $\alpha_{0,1}$, or $R_{\infty,1}$. Also for such (r_0,s_0), neither α_{r_0,s_0} nor $R_{1/r_0,1/s_0}$ dominates $\alpha_{r,s}$ or $R_{1/r,1/s}$ for any $(r,s) \notin Q(r_0,s_0)$.

(f) If $0<s_0<1$, then all of Remark (e) holds verbatim for $r_0=1$ (i.e. for the point $(r_0,s_0)=(1,s_0)$), except that α_{1,s_0} (and hence $R_{1,1/s_0}$) dominates $\alpha_{1,0}$ (and hence also $R_{1,\infty}$).

(g) If $0 < r_0 < 1$, then all of Remark (e) holds verbatim for $s_0 = 1$ (i.e. for the point $(r_0, s_0) = (r_0, 1)$), except that $\alpha_{r_0,1}$ (and hence $R_{1/r_0,1}$) dominates $\alpha_{0,1}$ (and hence also $R_{\infty,1}$).

The portion of this picture involving only $\alpha_{r,s}$, $r + s \le 1$, and $R_{p,q}$, $1/p + 1/q \le 1$, was summarized in [3, Theorem 4.1 and Remark 4.1]. (Some of that portion was already previously known.) The equation $\alpha_{1,1} = R_{1,1}$ is well known and elementary. What we have to do here in Section 1.2 is focus on the measures of dependence $\alpha_{r,s}$, $1 < r + s < 2$, and $R_{p,q}$, $1 < 1/p + 1/q < 2$. For these particular measures of dependence the (positive) domination results mentioned in Remark 1.2.1(e)(f)(g) follow immediately from interpolation theory. Since our (probability) terminology is different from the usual terminology of interpolation theory, we shall provide some of the details for the reader's convenience, in the form of Propositions 1.2.2-1.2.3 below. Then in Propositions 1.2.4-1.2.6 below, we shall give counterexamples to show that no other dominations occur for these measures of dependence.

Proposition 1.2.2. Suppose $0 \le r_0, s_0, r_1, s_1 \le 1$. Suppose that each of the following two statements (i) and (ii) holds:

(i) Either $r_0 = r_1 = 1$, or $r_0 \ne r_1$.

(ii) Either $s_0 = s_1 = 1$, or $s_0 \ne s_1$.

Suppose $0 < \theta < 1$. Define r and s by $r := (1-\theta)r_0 + \theta r_1$ and $s := (1-\theta)s_0 + \theta s_1$. Suppose that the equation $r + s \ge 1$ is satisfied. Then for some constant $C = C(r_0, r_1, s_0, s_1, \theta)$, one has that for every pair of σ-fields F and G,

$$R_{1/r,1/s}(F,G) \le C \cdot [\alpha_{r_0,s_0}(F,G)]^{1-\theta} \cdot [\alpha_{r_1,s_1}(F,G)]^{\theta}$$

In the case $r_0 \ne r_1$ and $s_0 \ne s_1$, Proposition 1.2.2 is an application of [19, Theorem 2.9] (and Remark 0.2). (The bilinear form B in eqn. (0.9) can be seen as a bilinear operator into $L_p(\Omega_1,F_1,P_1)$ where Ω_1 is trivial, consisting of just a single point; with this interpretation B "fits" the hypothesis of [19, Theorem 2.9].) In the case $r_0 = r_1 = 1$ and $s_0 \neq s_1$, one can take r_0 and r_1 out of the picture by an application of, say, [3, Lemma 3.7(i)], and then apply [19, Theorem 2.9] in an appropriate way to deal with s_0, s_1 and s. The other cases in Proposition 1.2.2 are then obvious.

<u>Proposition 1.2.3</u>.

(i) If $0 < r_0,s_0 < 1$ and $r_0 + s_0 > 1$, then α_{r_0,s_0} dominates $R_{1/r,1/s}$ for every $(r,s) \in Q(r_0,s_0) - \{(1,0),(r_0,s_0),(0,1)\}$.

(ii) If $0 < s_0 < 1$ then α_{1,s_0} dominates $R_{1/r,1/s}$ for every $(r,s) \in Q(1,s_0) - \{(1,s_0),(0,1)\}$.

(iii) If $0 < r_0 < 1$ then $\alpha_{r_0,1}$ dominates $R_{1/r,1/s}$ for every $(r,s) \in Q(r_0,1) - \{(1,0),(r_0,1)\}$.

<u>Proof</u>. The proofs of all three parts are similar, so we shall only prove (i). Let S_1 denote the line segment with endpoints (r_0,s_0) and $(1,0)$, and let S_2 denote the line segment with endpoints (r_0,s_0) and $(0,1)$. Recall that (trivially) $\alpha_{1,0} \le 1$ and $\alpha_{0,1} \le 1$. By Proposition 1.2.2, α_{r_0,s_0} dominates $R_{1/r,1/s}$ for every $(r,s) \in [S_1 \cup S_2] - \{(r_0,s_0),(1,0),(0,1)\}$. Also, trivially, α_{r_0,s_0} dominates $\alpha_{0,0}$ and hence also measures $R_{1/r,1/s}$, $r+s<1$ (which are equivalent to $\alpha_{0,0}$ as noted in [3, Remark 4.1]). Finally, the remaining points (r,s) in the interior of $Q(r_0,s_0)$ each lie on some line segment with one endpoint on $[S_1 \cup S_2] - \{(r_0,s_0),(1,0),(0,1)\}$ and the other in $\{(r,s): r+s<1\}$, and hence by the multilinear Thorin interpolation theorem (see e.g. [1, p. 18, Exercise 13]) α_{r_0,s_0} dominates $R_{1/r,1/s}$ for every such point (r,s).

Now we are ready for the construction of counterexamples. A few trivial facts are worth keeping in mind. The quantity $|P(A \cap B) - P(A)P(B)|$ remains unchanged if A is replaced by A^c, or B by B^c; consequently, one always has

$$\alpha_{r,s}(F,G) = \left[\begin{array}{l} \sup \dfrac{|P(A\cap B) - P(A)P(B)|}{[P(A)]^r[P(B)]^s} \\ A \in F,\ B \in G, \\ P(A) \le \tfrac{1}{2},\ P(B) \le \tfrac{1}{2} \end{array}\right]$$

Also, if F and G are finite σ-fields, each having exactly two atoms, then $R_{1/r,1/s}(F,G) \le 4\alpha_{r,s}(F,G)$ by a trivial argument.

Proposition 1.2.4. The following two statements hold:

(i) If $1 < p \le \infty$ then $R_{p,1}$ does not dominate $\alpha_{1,0}$.

(ii) If $1 < q \le \infty$ then $R_{1,q}$ does not dominate $\alpha_{0,1}$.

Proof. By symmetry it suffices to prove (i). For each ε, $0 < \varepsilon < \tfrac{1}{2}$, there exists a probability space and a pair of finite σ-fields $F = \{\Omega, A, A^c, \phi\}$ and $G = \{\Omega, B, B^c, \phi\}$ such that $P(A\cap B) = P(A) = \varepsilon$ and $P(B) = \tfrac{1}{2}$; and by a direct calculation, $R_{p,1}(F,G) \le 4\alpha_{1/p,1}(F,G) = 4\varepsilon^{1/p'}$ and $\alpha_{1,0}(F,G) = \tfrac{1}{2}$. Statement (i) follows, and this completes the proof.

Proposition 1.2.5. Suppose $0 < r_0, s_0, r, s \le 1$, $r_0 + s_0 > 1$, $r + s > 1$, and $(r,s) \notin Q(r_0,s_0)$. Then $R_{1/r_0,1/s_0}$ does not dominate $\alpha_{r,s}$.

Proof. Let $ax + by = c$ be an equation of a line containing (r_0,s_0) and one of the points $(0,1)$ or $(1,0)$, such that the points $(0,0)$ and (r,s) are in opposite half-planes determined by that line. By the assumptions in Proposition 1.2.5, we can (and do) take a,b, and c all positive. Thus $ar + bs > c$. Also (since $r_0 + s_0 > 1$) we have that $c = \max\{a,b\}$. Define $\varepsilon > 0$ by the equation $ar + bs = c + \varepsilon$.

For each n sufficiently large, there exists a probability space and a pair of finite σ-fields $F = \{\Omega, A, A^c, \phi\}$ and $G = \{\Omega, B, B^c, \phi\}$ such that $P(A) = n^{-a} \le \tfrac{1}{2}$, $P(B) = n^{-b} \le \tfrac{1}{2}$, and $P(A\cap B) = n^{-a-b} + n^{-c-\varepsilon}$. For such an n it can easily be checked that $R_{1/r_0,1/s_0}(F,G) \le 4\alpha_{1/r_0,1/s_0}(F,G) = 4n^{-\varepsilon}$

and that $\alpha_{r,s}(F,G) = 1$. Proposition 1.2.5 follows.

Proposition 1.2.6. Suppose that $0 < r,s \le 1$, $r + s > 1$, and either $r < 1$ or $s < 1$. Then $\alpha_{r,s}$ does not dominate $R_{1/r,1/s}$.

Proof. By symmetry, without loss of generality we can (and do) assume that $r < 1$. (So we allow the possibility $s = 1$.)

(In what follows, the symbols π, ϕ, and ψ do not take their usual meanings.)

Define the probability space (Ω, M, P) as follows: $\Omega := [0,1] \times \{0,1\}$ (the union of two disjoint intervals); M is the σ-field of Borel subsets of Ω, and P is defined by

$$P(A \times \{0\}) := \int_A \phi(x)dx \text{ and } P(A \times \{1\}) := \int_A (1 - \phi(x))dx \tag{1.2.1}$$

for every Borel subset $A \subset [0,1]$, where

$$\phi(x) := \begin{cases} \pi + \psi(x) & \text{if } x \in [0,\tfrac{1}{2}] \\ \pi - 2\cdot\int_0^{1/2}\psi(u)du & \text{if } x \in (\tfrac{1}{2},1] \end{cases} \tag{1.2.2}$$

where $0 < \pi < \tfrac{1}{2}$ and

$$\psi(x) := \begin{cases} \varepsilon x^{r-1} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

where

$$\varepsilon := \frac{\pi^s}{(\log 1/\pi)^{1-r}}$$

$$a := (\tfrac{1}{2})\pi^{s/(1-r)}$$

$$b := (\tfrac{1}{2})\pi^{(1-s)/r}$$

Note that by our assumption $r+s>1$ and simple arithmetic, we have that $0<a<b\le\frac{1}{2}$. Also, for π sufficiently small, one has that $2\cdot\int_0^{\frac{1}{2}}\psi(x)dx \le \pi$ and $\psi(x)\le\frac{1}{2}$ for all $x\in[0,\frac{1}{2}]$, and hence $0\le\phi(x)\le 1$ for all $x\in[0,1]$. Consequently, for π sufficiently small, eqn. (1.2.1) does indeed define a probability measure. (We restrict π to such small values.)

Define the "marginal" σ-fields F and G by

$$F := \{A\times\{0,1\}:\ A\subset[0,1]\text{ Borel set}\}$$

$$G := \{[0,1]\times B:\ B = \{0,1\},\ \{0\},\ \{1\},\ \phi\}$$

Define the event $B_0 := [0,1]\times\{0\}$. Note that the marginal of P on [0,1] is Lebesgue measure, and that $P(B_0) = \pi < \frac{1}{2}$.

We shall first get an upper bound on $\alpha_{r,s}(F,G)$. First a preliminary calculation will be handy. The function εx^{r-1} is non-increasing on $(0,\infty)$. Also, $\phi(x)-\pi$ is non-negative if $0\le x\le\frac{1}{2}$, and negative if instead $\frac{1}{2}<x\le 1$. Hence, letting m denote Lebesgue measure, we have that for every Borel subset $A\subset[0,1]$,

$$\begin{aligned}\int_A[\phi(x)-\pi]dx &\le \int_{A\cap[0,\frac{1}{2}]}\psi(x)dx \le \int_{A\cap[0,\frac{1}{2}]}\varepsilon x^{r-1}dx\\ &\le \int_0^{m(A\cap[0,\frac{1}{2}])}\varepsilon x^{r-1}dx\\ &= (\varepsilon/r)\cdot[m(A\cap[0,\tfrac{1}{2}])]^r \le (\varepsilon/r)\cdot[m(A)]^r\end{aligned}$$

and

$$\begin{aligned}\int_A[\phi(x)-\pi]dx &\ge -m(A\cap(\tfrac{1}{2},1])\cdot 2\cdot\int_0^{\frac{1}{2}}\psi(x)dx\\ &\ge -m(A\cap(\tfrac{1}{2},1])\cdot 2\cdot\int_0^{\frac{1}{2}}\varepsilon x^{r-1}dx\\ &= -(\varepsilon/r)\cdot m(A\cap(\tfrac{1}{2},1])\cdot(\tfrac{1}{2})^{r-1}\\ &\ge -(\varepsilon/r)\cdot m(A\cap(\tfrac{1}{2},1])\cdot[m(A\cap(\tfrac{1}{2},1])]^{r-1}\\ &\ge -(\varepsilon/r)\cdot[m(A)]^r\end{aligned}$$

and hence $|\int_A[\phi(x) - \pi]dx| \le (\varepsilon/r)\cdot[m(A)]^r$.

Consequently, it is easy to see that

$$\alpha_{r,s}(F,G) = \sup_{A_0 \in F} \frac{|P(A_0 \cap B_0) - P(A_0)P(B_0)|}{[P(A_0)]^r \cdot [P(B_0)]^s}$$

$$= \sup_{A \subset [0,1] \text{ Borel set}} \frac{|\int_A \phi(x)dx - [m(A)]\cdot\pi|}{[m(A)]^r \cdot \pi^s}$$

$$\le \sup_{A \subset [0,1] \text{ Borel set}} \frac{(\varepsilon/r)\cdot[m(A)]^r}{[m(A)]^r \cdot \pi^s}$$

$$= \frac{1}{r\cdot(\log 1/\pi)^{1-r}}$$

Note that $\alpha_{r,s}(F,G)$ becomes arbitrarily small for π sufficiently small. (We are using our assumption $r < 1$ here.)

Now we only need to show that $R_{1/r,1/s}(F,G)$ fails to become small with π. By Proposition 0.1 and symmetry, we have for $p := 1/r$ and $q := 1/s$ (so that $1/p' = 1 - r$),

$$R_{p,q}(F,G) = \sup_{g \text{ complex-valued, simple, } G\text{-measurable}} \frac{\|E(g|F) - Eg\|_{p'}}{\|g\|_q}$$

$$\ge \frac{\|E(I_{B_0}|F) - EI_{B_0}\|_{p'}}{\|I_{B_0}\|_q}$$

$$= \frac{[\int_0^1 |\phi(x) - \pi|^{p'} dx]^{1/p'}}{\pi^s}$$

$$\geq \frac{[\int_a^b (\psi(x))^{p'} dx]^{1/p'}}{\pi^s}$$

$$= \frac{[\int_a^b (\varepsilon \cdot x^{r-1})^{p'} dx]^{1-r}}{\pi^s}$$

$$= \frac{\varepsilon \cdot [\int_a^b x^{-1} dx]^{1-r}}{\pi^s}$$

$$= \frac{1}{(\log 1/\pi)^{1-r}} \; (\log b/a)^{1-r}$$

$$= \frac{1}{(\log 1/\pi)^{1-r}} \cdot [(\frac{1-s}{r} - \frac{s}{1-r})(\log \pi)]^{1-r}$$

$$= (\frac{s}{1-r} - \frac{1-s}{r})^{1-r}$$

Since $r + s > 1$, it follows from elementary arithmetic that $\frac{s}{1-r} - \frac{1-s}{r}$ is a positive constant. Hence $R_{p,q}(F,G)$ fails to converge to 0 as $\pi \to 0$. This completes the proof.

SECTION 1.3: APPLICATION OF THE REITERATION PROCEDURE

Here we give another proof of [3, Theorem 2.2]. The proof is rather simple but requires more interpolation theory, and seems to be harder to generalize to the multidimensional case as in [3, Theorem 2.1]. (Such a generalization of the proof might be possible if one uses as a tool a version of Zafran's [19, Theorem 2.9] multilinear Marcinkiewicz interpolation theorem with an explicit upper bound on the constant.) We shall restate [3, Theorem 2.2] in a somewhat loose form (applicable to some other contexts besides probability spaces); the main emphasis is on the proof.

Theorem 1.3.1. Suppose T is a linear operator, $1 < p < \infty$, and

(1.3.1) $T: L_1 \to L_1$ with norm ≤ 1,

(1.3.2) $T: L_\infty \to L_\infty$ with norm ≤ 1,

(1.3.3) $T: L_p \to L_{p,\infty}$ with norm $\le \varepsilon \le 1$.

Then

(1.3.4) $T: L_p \to L_p$ with norm $\le C\varepsilon(1 + \log 1/\varepsilon)^{1/p}$ where C depends only on p.

Note that eqns. (1.3.1) and (1.3.2) imply (1.3.3) and (1.3.4) for $\varepsilon = 1$ by standard interpolation theorems (see [1]). The point of Theorem 1.3.1 is that a small norm in eqn. (1.3.3) forces the norm in eqn. (1.3.4) to be almost as small, i.e. within a log term. (For the definition of the Lorentz space $L_{p,\infty}$, see e.g. [1] or [19].)

Proof. Throughout this proof, the letter C is used only for multiplicative constants which depend only on p. The value of C may vary from one appearance to the next.

The trivial cases $\varepsilon = 1$ or 0 can be omitted. Moreover, to cover the remaining cases, without loss of generality we can restrict ε to small values. Henceforth, we impose the condition $\varepsilon \le e^{-2}$.

Let $\delta = -1/(\log \varepsilon)$. Then $0 < \delta \le \frac{1}{2}$. Define p_0 and p_1 by $1/p_0 = (1 - \delta)/p + \delta/1$ and $1/p_1 = (1 - \delta)/p + \delta/\infty$. Then $1 < p_0 < p < p_1 < \infty$. Now we apply the Marcinkiewicz interpolation theorem twice, each time with an explicit upper bound on the constant in that theorem; see e.g. [20, Chapter 12, Theorem (4.6) and eqn. (4.2.1)]. In that way, by eqns. (1.3.1) and (1.3.3) we obtain $T: L_{p_0} \to L_{p_0}$ with norm $\le C(1/\delta)^{1/p_0}\varepsilon^{1-\delta}$ and by eqns. (1.3.2) and (1.3.3) we obtain $T: L_{p_1} \to L_{p_1}$ with norm $\le C(1/\delta)^{1/p_1}\varepsilon^{1-\delta}$. Let $\theta = 1 - 1/p$. Then $0 < \theta < 1$ and $1/p = (1 - \theta)/p_0 + \theta/p_1$. By applying the Riesz (or, in the complex case, Riesz-Thorin) interpolation theorem (see [1]) we have that $T: L_p \to L_p$

with $\text{norm} \le C\cdot(1/\delta)^{1/p}\varepsilon^{1-\delta} = C(-\log\varepsilon)^{1/p}\cdot\varepsilon\cdot\varepsilon^{-\delta} \le C(-\log\varepsilon)^{1/p}\cdot\varepsilon$ (since $\varepsilon^{-\delta} = e$). This completes the proof.

SECTION 1.4: A SHARP DOMINATION INEQUALITY

Here our aim is to find a sharp upper bound on $R_{p,p'}$ based on $\alpha_{1/p,1/p'}$ in some cases where one of the two σ-fields has two atoms. This complements Section 4.4 of [3].

Theorem 1.4.1. Suppose F and G are σ-fields, with G being of the form $G = \{\Omega, B_0, B_1, \phi\}$ where $P(B_0) = \frac{1}{2}$ and $B_1 = B_0^c$. If $1 < p < \infty$ then

$$R_{p,p'}(F,G) \le 2\alpha_{1/p,1/p'}(F,G)\cdot[1 - (\tfrac{1}{p})^{p'}p' \log(2\alpha_{1/p,1/p'}(F,G))]^{1/p'}.$$

Further, this inequality is sharp; for each $\alpha \in [0,\frac{1}{2}]$ there exists a probability space with σ-fields F and $G = \{\Omega, B_0, B_1, \phi\}$ where $P(B_0) = \frac{1}{2}$ and $B_1 = B_0^c$, such that $\alpha_{1/p,1/p'}(F,G) = \alpha$ and $R_{p,p'}(F,G) = 2\alpha[1 - (\frac{1}{p})^{p'}p' \log(2\alpha)]^{1/p'}$.

In particular, referring to eqns. (0.6) and (0.8),

$$\rho(F,G) \le 2\lambda(F,G)\cdot[1 - \tfrac{1}{2}\log(2\lambda(F,G))]^{\frac{1}{2}}.$$

Under the hypothesis of Theorem 1.4.1, one automatically has that $\alpha_{1/p,1/p'}(F,G) \le \frac{1}{2}$. By the comments prior to Proposition 1.2.4 in Section 1.2, in evaluating $\alpha_{1/p,1/p'}(F,G)$ one only needs to consider events A and B with probability $\le \frac{1}{2}$. Since $0 \le P(A\cap B_0) \le P(A)$, one has that if $P(A) \le \frac{1}{2}$ then

$$|P(A\cap B_0) - P(A)P(B_0)| \le (\tfrac{1}{2})P(A) \le (\tfrac{1}{2})[P(A)]^{1/p}[P(B_0)]^{1/p'}$$

(where B_0 is as in Theorem 1.4.1). Consequently $\alpha_{1/p,1/p'}(F,G)$ cannot exceed $\frac{1}{2}$ under these circumstances.

The proof of Theorem 1.4.1 will be given at the end of this section and will be based on some preliminary lemmas.

We start from the purely discrete case, when F (as well as G of course) is finitely generated. The reasoning is close to that in [12]. We use the extreme point method and the following version of the Krein-Milman theorem:

<u>Theorem A</u>. If $K \in \mathbb{R}^n$ is a compact convex set, then for each $x \in K$ there exist extreme points $x_1, \ldots, x_{n+1}$ of K and non-negative numbers $\lambda_1, \ldots, \lambda_{n+1}$ such that $\Sigma_{i=1}^{n+1} \lambda_i = 1$ and $x = \Sigma_{i=1}^{n+1} \lambda_i x_i$.

(Recall that $x \in K$ is called an extreme point of K if x cannot be represented as a non-trivial convex combination of points of K.)

We shall be interested in a particular compact convex set of probability measures on a certain measurable space. Let $\alpha \in [0, \frac{1}{2}]$ and $r, s \in [0,1]$ be fixed, and suppose m is a fixed positive *even* integer. Define the sets $\Omega_1 := \{1, 2, \ldots, m\}$ and $\Omega_2 = \{0,1\}$. The measurable space will be $\Omega_1 \times \Omega_2$ (accompanied by the discrete σ-field). For ease of notation we shall denote $A_i := \{i\}$, $i = 1, 2, \ldots, m$, and $B_0 := \{0\}$, $B_1 := \{1\}$. Let K be the set of all probability measures on $\Omega_1 \times \Omega_2$ such that

(1.4.1) $\quad P(A_i \times \Omega_2) = 1/m$ for all $i = 1, \ldots, m$,

(1.4.2) $\quad P(\Omega_1 \times B_0) = P(\Omega_1 \times B_1) = \frac{1}{2}$, and

(1.4.3) $\quad |P(A \times B) - P(A \times \Omega_2) P(\Omega_1 \times B)| \leq \alpha \cdot [P(A \times \Omega_2)]^r \cdot [P(\Omega_1 \times B)]^s$

for each pair of subsets $A \subset \{1, \ldots, m\}$, $B \subset \{0,1\}$.

Clearly K can be regarded as a compact subset of $\mathbb{R}^{m-1}$. Note that K is convex too. If $P_1, P_2 \in K$ and $0 < \lambda < 1$, then the probability measure $P := (1-\lambda)P_1 + \lambda P_2$ satisfies eqns. (1.4.1) and (1.4.2), and since the marginals of P_1 and P_2 (and hence also P) are equal, eqn. (1.4.3) is also satisfied by P, and hence $P \in K$.

We shall be interested in the σ-fields

$$
(1.4.4) \qquad \begin{aligned} F_m &:= \{A \times \Omega_2 : A \subset \{1,2,\ldots, m\}\} \text{ and} \\ G_m &:= \{\Omega_1 \times B : B \subset \{0,1\}\}. \end{aligned}
$$

(The subscript m serves as a reminder of the parameter m on which $\Omega_1 \times \Omega_2$ is based.) K is the set of all probability measures on $\Omega_1 \times \Omega_2$ with uniform marginals such that $\alpha_{r,s}(F_m, G_m) \le \alpha$.

Lemma 1.4.2. Each extreme point $P \in K$ has the form

$$P(A_{\sigma(i)} \times B_1) = 1/(2m) + F(i/m) - F((i-1)/m) \quad \text{for} \quad 1 \le i \le m/2,$$

$$P(A_{\sigma(i)} \times B_1) = 1/m - P(A_{\sigma(m+1-i)} \times B) \quad \text{for} \quad m/2 + 1 \le i \le m,$$

where σ is a permutation of $\{1,2,\ldots, m\}$ and $F(x) := \min\{x/2, \alpha \cdot 2^{-s} x^r\}$.

Sketch of proof. Each $P \in K$ is uniquely determined by the numbers $\Delta_i := P(A_i \times B_1) - 1/(2m)$, $i = 1,2,\ldots, m$. By eqns. (1.4.1) and (1.4.2) these numbers satisfy

$$(1.4.5) \qquad \sum_{i=1}^{m} \Delta_i = 0 \quad \text{and}$$

$$(1.4.6) \qquad |\Delta_i| \le (2m)^{-1} \quad \text{for all } i = 1,\ldots, m.$$

Since the property of being an extreme point of K is not affected by the order of numeration of events $A_1, A_2, \ldots, A_m$, to prove Lemma 1.4.2 we may assume without loss of generality that $\Delta_1 \ge \Delta_2 \ge \ldots \ge \Delta_m$. By eqns. (1.4.3) and (1.4.6) we see that a given sequence of numbers $\Delta_1 \ge \Delta_2 \ge \ldots \ge \Delta_m$ corresponds to a probability $P \in K$ if and only if (1.4.5) and the following two equations hold:

$$(1.4.7) \qquad \text{For all } k = 1,\ldots, m, \quad \sum_{i=1}^{k} \Delta_i \le \min\{\frac{k}{2m}, \alpha \cdot (\frac{k}{m})^r \cdot 2^{-s}\},$$

$$(1.4.8) \qquad \text{For all } k = 1,\ldots, m, \quad \sum_{i=m+1-k}^{m} (-\Delta_i) \le \min\{\frac{k}{2m}, \alpha \cdot (\frac{k}{m})^r \cdot 2^{-s}\}.$$

Using this information, one can check with a little elementary work that the unique extreme point P of K which satisfies $\Delta_1 \geq \Delta_2 \geq \ldots \geq \Delta_m$ is the one in which equality in (1.4.7) and (1.4.8) is attained for all $k = 1, \ldots, m/2$, i.e.

$$(1.4.9) \quad \begin{aligned} \Delta_i &= F(i/m) - F((i-1)/m) \quad \text{for all } i \leq m/2, \\ \Delta_i &= -\Delta_{m+1-i} \quad \text{for all } i \geq m/2+1, \end{aligned}$$

where F is as in the statement of Lemma 1.4.2. The lemma follows.

Remark. Referring to the proof of Lemma 1.4.2, in the case where $r + s = 1$, the extreme point $P \in K$ given by eqn. (1.4.9) has the property that for $A = \{1, \ldots, m/2\}$, $P(A \cap B_1) - P(A)P(B_1) = \alpha \cdot [P(A)]^r [P(B_1)]^s$ and hence (for this P) $\alpha_{r,s}(F_m, G_m) = \alpha$. (See the paragraph preceding Lemma 1.4.2.) The same holds for any other extreme point of K with an appropriate modification of the event A.

Lemma 1.4.3. Let $\Delta_1, \ldots, \Delta_m$ now be fixed as in eqn. (1.4.9). If $1 < p, q < \infty$, then for any probability $P \in K$ we have that

$$R_{p,q}(F_m, G_m) \leq 2\left(\sum_{i=1}^{m} |\Delta_i|^{p'}\right)^{1/p'} m^{1/p}$$

(where F_m and G_m are as in eqn. (1.4.4)). Moreover, equality is attained for any extreme point P of K.

Proof. By Theorem A, each probability $P \in K$ can be represented as a finite convex combination of extreme points of K, i.e. $P = \Sigma_\sigma \lambda_\sigma P_\sigma$ where $\Sigma_\sigma \lambda_\sigma = 1$, $\lambda_\sigma \geq 0$, and the P_σ are the extreme points, given by Lemma 1.4.2. Here σ is ranging over all permutations of $\{1, 2, \ldots, m\}$.

Since the probability measures in K all have the same marginals, we have that for any $P \in K$ and any r.v.'s X and Y on $\Omega_1 \times \Omega_2$ such that X is F_m-measurable (i.e. a function only of the first coordinate) and Y is G_m-measurable (i.e. a function only of the second coordinate),

$$\left|\int XYdP - \int XdP\int YdP\right| \le \sum_{\sigma} \lambda_\sigma \left|\int XYdP_\sigma - \int XdP_\sigma\int YdP_\sigma\right|$$

$$\le \max_\sigma \left|\int XYdP_\sigma - \int XdP_\sigma\int YdP_\sigma\right|$$

$$\le a\cdot\left(\int |X|^p dP_\sigma\right)^{1/p}\cdot\left(\int |Y|^q dP_\sigma\right)^{1/q}$$

where $a := \sup_{X,Y}\left|\int XYdP_\sigma - \int XdP_\sigma\int YdP_\sigma\right|/(\|X\|_p \cdot \|Y\|_q)$

(the sup being taken over X and Y as above). Note that a depends on m but is not dependent on σ.

If X and Y are as above and we add a constant to Y, then $\left|\int XYdP_\sigma - \int XdP_\sigma\int YdP_\sigma\right|$ remains unchanged; but $\|Y\|_q$ can be minimized (under addition of a constant) by centering it at expectation. This latter fact follows from the nature of the σ-field G_m (i.e. two atoms, each with probability ½). Consequently,

$$a = \sup_X 2\left|\int XYdP_\sigma\right|$$

where $Y := I_{B_1} - \frac{1}{2}$ and the sup is taken over all F_m-measurable X such that $\|X\|_p = 1$.

From the representation $X = \Sigma_{i=1}^m x_i I_{A_i}$ we see that a is the maximum of the function $f(x_1,\ldots, x_m) := 2\Sigma_{i=1}^m x_i\Delta_i$ under the restriction $(\Sigma_{i=1}^m |x_i|^p)^{1/p} = m^{1/p}$. It is well known that this maximum is $2m^{1/p}(\Sigma_{i=1}^m |\Delta_i|^{p'})^{1/p'}$. Lemma 1.4.3 (both parts) is now easy to see.

Lemma 1.4.4. Suppose that $0 \le r,s \le 1$ and $1 \le p,q \le \infty$. Suppose that F and G are σ-fields on some probability space, where G has two atoms B and B^c, each having probability ½. Then there exists a sequence $P_2, P_4, P_6, \ldots$ of probability measures, where for each (even) m the measure P_m is defined on $\{1,2,\ldots, m\} \times \{0,1\}$ and has uniform marginals, such that

$$\alpha_{r,s}(F,G) = \lim_{m\to\infty} \alpha_{r,s}(F_m,G_m)$$

$$R_{p,q}(F,G) = \lim_{m\to\infty} R_{p,q}(F_m,G_m)$$

where for each (even) m the σ-fields F_m and G_m are as in eqn. (1.4.4) and $\alpha_{r,s}(F_m,G_m)$ and $R_{p,q}(F_m,G_m)$ are as defined with respect to the probability measure P_m. (In the two limits, m is restricted to even integers.)

Proof. By an elementary argument based on eqn. (0.1), one can show that it suffices to prove this lemma in the case where F is itself a finite σ-field. We shall henceforth make that assumption on F. Enlarging the given probability space if necessary we may assume without loss of generality that there exists a r.v. U which is independent of $F \vee G$ and has a continuous distribution function. By an elementary argument,

$$\text{(1.4.10)} \quad \begin{aligned} \alpha_{r,s}(F \vee \sigma(U),G) &= \alpha_{r,s}(F,G) \quad \text{and} \\ R_{p,q}(F \vee \sigma(U),G) &= R_{p,q}(F,G) \end{aligned}$$

(where $\sigma(\ldots)$ denotes the σ-field generated by $(\ldots)$).

Let Z be a $F \vee \sigma(U)$-measurable r.v. with a continuous distribution function $F_Z(\cdot)$, such that $F \subset \sigma(Z)$. (For example, take $Z = \Sigma_{i=1}^{n} I_{A(i)} \cdot f_i(U)$ where $A(1),\ldots, A(n)$ are the atoms of F and $f_1,\ldots, f_n$ are suitably chosen functions.) Then as a consequence of eqn. (1.4.10) we have that

$$\alpha_{r,s}(\sigma(Z),G) = \alpha_{r,s}(F,G) \quad \text{and}$$

$$R_{p,q}(\sigma(Z),G) = R_{p,q}(F,G).$$

Now consider the probability measure P on $[0,1] \times \{0,1\}$ induced by the random vector $(F_Z(Z), I_B)$ (where B is one of the atoms of G). Then P has uniform marginals, and defining the σ-fields F^* and G^* generated by the first and second coordinates of this new probability space, we have that

$$\alpha_{r,s}(F^*,G^*) = \alpha_{r,s}(\sigma(Z),G) = \alpha_{r,s}(F,G) \quad \text{and}$$

$$R_{p,q}(F^*,G^*) = R_{p,q}(\sigma(Z),G) = R_{p,q}(F,G).$$

For each even m let F_m^* denote the sub-σ-field of F^* generated by the events $\{((i-1)/m, i/m) \times \{0,1\},\ i = 1,2,\ldots, m\}$. Then

$$\lim_{m\to\infty} \alpha_{r,s}(F_m^*,G^*) = \alpha_{r,s}(F^*,G^*) = \alpha_{r,s}(F,G) \quad \text{and}$$

(1.4.11)

$$\lim_{m\to\infty} R_{p,q}(F_m^*,G^*) = R_{p,q}(F^*,G^*) = R_{p,q}(F,G).$$

Now if for each even m we define the probability measure P_m on $\{1,2,\ldots, m\} \times \{0,1\}$ by $P_m(\{i\} \times \{j\}) = P(((i-1)/m, i/m) \times \{0,1\})$, $i = 1,2,\ldots,m$, $j = 0,1$, then it is easy to see that as a trivial corollary of eqn. (1.4.11) these measures P_m satisfy the conclusion of Lemma 1.4.4.

Proof of Theorem 1.4.1. To prove the inequality in Theorem 1.4.1, by Lemmas 1.4.3 and 1.4.4 it suffices to notice that if $0 < r < 1$, $s = 1 - r$, $p = 1/r$ and hence $p' = 1/s$, then

$$\lim_{m\to\infty} 2m^{1/p}\left(\sum_{i=1}^{m} |\Delta_i|^{p'}\right)^{1/p'} = 2\alpha \cdot \left[1 - \left(\frac{1}{p}\right)^{p'} p' \log(2\alpha)\right]^{1/p'}$$

(where $\Delta_1, \ldots, \Delta_m$ are as in eqn. (1.4.9)).

To prove the second part of Theorem 1.4.1 consider the probability measure P on $[0,1] \times \{0,1\}$ with uniform marginals, such that for every Borel subset A of $[0,1]$ (letting $B_1 = \{1\}$ as above),

$$P(A \cap B_1) = \int_A [(\tfrac{1}{2}) + F'(x)]dx$$

where $F(x)$ is as in Lemma 1.4.2 for $0 \le x \le \frac{1}{2}$ and $F(x) = F(1-x)$ for $\frac{1}{2} \le x \le 1$. By an argument like the last part of the proof of Lemma 1.4.4 (and taking into account the Remark prior to Lemma 1.4.3) one can verify that if $0 < r < 1$,

$s = 1 - r$, $p = 1/r$ and hence $p' = 1/s$, then for this measure P, $\alpha_{1/p,1/p'}(F^*,G^*) = \alpha$ and $R_{p,p'}(F^*,G^*) = 2\alpha[1 - (\frac{1}{p})^{p'} p' \log(2\alpha)]^{1/p'}$ where F^* and G^* are as in the proof of Lemma 1.4.4.

SECTION 2: THE VECTOR CASE

In Section 2 here we shall examine measures of dependence which are based on random variables taking their values in Banach spaces, i.e. "B-valued" r.v.'s. For simplicity we consider only real Banach spaces. But all results here are valid for complex Banach spaces as well. In fact every complex Banach space is also a real Banach space (by restricting the scalars), and a simple argument will show that the results for the complex case follow from the real case.

Much of the theory of operators on spaces of functions (including many results in interpolation theory) has been extended to functions taking their values in Banach spaces. Most of the results here in Section 2 are entirely elementary and in many cases are simply a reformulation, in our context, of results already known in functional analysis in connection with B-valued functions. Our main goal here is to clarify the connection between measures of dependence and functional analysis involving B-valued functions.

SECTION 2.1: PRELIMINARIES

Here in Section 2.1 we shall review, in our terminology involving measures of dependence, some elementary aspects of functional analysis involving B-valued functions.

The norm of an element x of a given Banach space B will be denoted by $\|x\|_B$.

First recall that in the special case of a real Hilbert space H with inner product $\langle .,. \rangle$, each real bounded linear functional g on H is of the form $g(x) := \langle x,y \rangle \ \forall x \in H$ for some fixed $y \in H$. In this setting it is common practice to identify the functional g with the element y; and with this identification one can say that $g(x) = \langle x,g \rangle$. To put this another way, H is (isometric to) its own (real) dual space.

Suppose B is a Banach space (with just real scalar multiplication). If $x \in B$ and g is a real bounded linear functional on B, then we shall use the notation $\langle x,g \rangle := g(x)$ (by analogy with the Hilbert space context described above). The (real) dual space of B, i.e. the space of all real bounded linear functionals on B, will be denoted by B*. Of course B* is itself a Banach space (with just real scalar multiplication) with norm $\|\cdot\|_{B^*}$ given by

(2.1.1) $$\|g\|_{B^*} := \sup \frac{|\langle x,g \rangle|}{\|x\|_B}, \quad x \in B$$

$$= \sup \frac{\langle x,g \rangle}{\|x\|_B}, \quad x \in B$$

(The latter equality is trivial; one can consider -x in place of x.)

If $x \in B$ then we can define a real linear functional z on B* by $z(y) := \langle x,y \rangle \ \forall y \in B^*$. That is, $\langle y,z \rangle := \langle x,y \rangle \ \forall y \in B^*$. It is easy to see that z is a bounded functional, i.e. $z \in B^{**}$ (the "second dual" of B), with norm $\|z\|_{B^{**}} = \|x\|_B$. The identification of each $x \in B$ with the corresponding $z \in B^{**}$ is the standard way of isometrically embedding B in B**. In many cases B** will have other elements besides (the embedding of) the ones in B.

Recall the simple equation

(2.1.2) $$\forall y \in B^*, \quad \left[\sup \frac{|\langle y,z \rangle|}{\|z\|_{B^{**}}}, \quad z \in B^{**}\right] = \|y\|_{B^*}$$

(or, similarly, $\forall x \in B$, $\|x\|_B = \sup |\langle x,y \rangle| / \|y\|_{B^*}$, $y \in B^*$) . The following simple consequence of eqn. (2.1.2) will be useful later on:

Lemma 2.1.1: Suppose B is a Banach space, $y \in B^*$, $z \in B^{**}$, and $\varepsilon > 0$. Then there exists $x \in B$ with $\|x\|_B \le \|z\|_{B^{**}}$ such that $|\langle y,z\rangle| \le (1+\varepsilon) \cdot \langle x,y\rangle$.

Now let us return to our probability space (Ω,M,P) . For each σ-field $A \subset M$ and each Banach space B, define $S(A,B)$ to be the set of all A-measurable simple B-valued random variables, that is, the set of random variables of the form $\Sigma_{i=1}^N x_i I_{A(i)}$ where N is a positive integer, $x_1,\ldots,x_N$ are elements of B, and $\{A_1,\ldots,A_N\}$ is a partition of Ω with $A_i \in A\ \forall i$. For each $X \in S(A,B)$ and each p, $1 \le p \le \infty$, define

$$\|X\|_p := \| \ \|X\|_B \ \|_p := \begin{cases} [E(\|X\|_B^p)]^{1/p} & \text{if } 1 \le p < \infty \\ \text{ess sup } \|X\|_B & \text{if } p = \infty \end{cases}$$

To put this another way, if $X = \Sigma_{i=1}^N x_i I_{A(i)}$ with $N, x_1,\ldots,x_N, A_1,\ldots,A_N$ as above, then $\|X\|_p = [\Sigma_{i=1}^N \|x_i\|_B^p \cdot P(A_i)]^{1/p}$ if $1 \le p < \infty$, and $\|X\|_\infty = \max\{\|x_i\|_B : P(A_i) > 0\}$.

For each p,q with $1 \le p,q \le \infty$ and each Banach space B, define the measure of dependence $R_{p,q}(B)$ between σ-fields F and G as follows:

$$R_{p,q}(F,G;B) := \sup \frac{|E\langle X,Y\rangle - \langle EX,EY\rangle|}{\|X\|_p \cdot \|Y\|_q} , \quad \begin{matrix} X \in S(F,B) \\ Y \in S(G,B^*) \end{matrix}$$

Proposition 2.1.2: Suppose $1 \le p,q \le \infty$, and B is a Banach space. Then for any two σ-fields F and G the following three statements hold:

(a) For any Banach space B_1 which is isometric to B, $R_{p,q}(F,G;B_1) = R_{p,q}(F,G;B)$.

(b) For any subspace B_1 of B (or Banach space B_1 which is isometric to a subspace of B), $R_{p,q}(F,G;B_1) \leq R_{p,q}(F,G;B)$.

(c) $R_{p,q}(F,G;B) = \begin{bmatrix} \sup R_{p,q}(F,G;B_1) \\ \text{Finite-dimensional subspaces } B_1 \text{ of } B \end{bmatrix}$

Proposition 2.1.2 is elementary and its proof is left to the reader. The Hahn-Banach theorem will be helpful in proving part (b).

<u>Proposition 2.1.3</u>: If $1 \leq p,q \leq \infty$ and B is a Banach space, then for any two σ-fields F and G, $R_{p,q}(F,G;B^*) = R_{q,p}(G,F;B)$.

<u>Proof</u>: We shall first prove the following lemma:

<u>Lemma</u>: If $\varepsilon > 0$, $Y \in S(F,B^*)$, and $Z \in S(G,B^{**})$, then there exists $X \in S(G,B)$ such that $\|X\|_q \leq \|Z\|_q$ and $|E\langle Y-EY,Z\rangle| \leq (1+\varepsilon)\, E\langle X,Y-EY\rangle$.

<u>Proof of Lemma</u>: Represent $Y-EY$ and Z by $Y-EY = \Sigma_{i=1}^{I} y_i I_{A(i)}$ and $Z = \Sigma_{j=1}^{J} z_j I_{B(j)}$ where $y_i \in B^* \;\forall i$, $\{A_1,\ldots,A_I\}$ is a partition of Ω with $A_i \in F \;\forall i$, $z_j \in B^{**} \;\forall j$, and $\{B_1,\ldots,B_J\}$ is a partition of Ω with $B_j \in G \;\forall j$.

Using Lemma 2.1.1, for each $j=1,\ldots,J$ let $x_j \in B$ be such that $\|x_j\|_B \leq \|z_j\|_{B^{**}}$ and

$$|\langle \Sigma_{i=1}^{I} P(A_i \cap B_j) y_i \, , \, z_j \rangle| \leq (1+\varepsilon) \langle x_j \, , \, \Sigma_{i=1}^{I} P(A_i \cap B_j) y_i \rangle \quad .$$

Define $X \in S(G,B)$ by $X := \Sigma_{j=1}^{J} x_j I_{B(j)}$. Then $\|X(\omega)\|_B \leq \|Z(\omega)\|_{B^{**}}$ for every $\omega \in \Omega$, and hence $\|X\|_q \leq \|Z\|_q$. Also,

$$\begin{aligned} |E\langle Y-EY,Z\rangle| &= |\Sigma_{j=1}^{J} \langle \Sigma_{i=1}^{I} P(A_i \cap B_j) y_i, \, z_j \rangle| \\ &\leq \Sigma_{j=1}^{J} |\langle \Sigma_{i=1}^{I} P(A_i \cap B_j) y_i, \, z_j \rangle| \\ &\leq \Sigma_{j=1}^{J} (1+\varepsilon) \cdot \langle x_j, \, \Sigma_{i=1}^{I} P(A_i \cap B_j) y_i \rangle \\ &= (1+\varepsilon)\, E\langle X,Y-EY\rangle \end{aligned}$$

This completes the proof of the lemma.

By the lemma, for any $Y \in S(F,B^*)$, $Z \in S(G,B^{**})$, and $\varepsilon > 0$,

$$\frac{|E\langle Y,Z\rangle - \langle EY,EZ\rangle|}{\|Y\|_p \ \|Z\|_q} = \frac{|E\langle Y-EY,Z\rangle|}{\|Y\|_p \ \|Z\|_q}$$

$$\leq (1+\varepsilon) \cdot \left[\sup \frac{E\langle X,Y-EY\rangle}{\|X\|_q \ \|Y\|_p} \ , \quad X \in S(G,B)\right]$$

$$= (1+\varepsilon) \cdot \left[\sup \frac{E\langle X,Y\rangle - \langle EX,EY\rangle}{\|X\|_q \ \|Y\|_p} \ , \quad X \in S(G,B)\right]$$

$$\leq (1+\varepsilon) \cdot R_{q,p}(G,F;B)$$

Hence $R_{p,q}(F,G;B^*) \leq R_{q,p}(G,F;B)$.

By an exactly analogous argument, $R_{q,p}(G,F;B^{**}) \leq R_{p,q}(F,G;B^*)$. By Proposition 2.1.2(b) and the fact (noted earlier) that B is isometric to a subspace of B^{**}, $R_{q,p}(G,F;B) \leq R_{q,p}(G,F;B^{**})$. From these last three inequalities Proposition 2.1.3 follows. This completes the proof.

From two applications of Proposition 2.1.3 (or from its proof) we have

<u>Corollary 2.1.4</u>: If $1 \leq p,q \leq \infty$ and B is a Banach space, then $R_{p,q}(B^{**}) = R_{p,q}(B)$.

Let ℓ^1 (resp. ℓ^∞) denote the Banach space of all absolutely summable (resp. bounded) sequences $x := (\xi_1,\xi_2,\xi_3,\ldots)$ of real numbers, with norm

$$\|x\|_{\ell^1} := \Sigma_{i=1}^{\infty} |\xi_i| \quad (\text{resp.} \quad \|x\|_{\ell^\infty} := \sup_i |\xi_i|) .$$

It is well known that $(\ell^1)^*$ is isometric to ℓ^∞, the standard isometry being (informally described) as follows: If $x := (\xi_1,\xi_2,\ldots) \in \ell^1$ and $y := (\eta_1,\eta_2,\ldots) \in \ell^\infty$ then

$$(2.1.3)\qquad y(x) = \langle x,y\rangle := \Sigma_{i=1}^{\infty} \xi_i \eta_i$$

Theorem 2.1.5: If $1 \leq p,q \leq \infty$ and B is a non-trivial Banach space, then $R_{p,q}(\mathbb{R}) \leq R_{p,q}(B) \leq R_{p,q}(\ell^{\infty}) = R_{p,q}(\ell^1)$.

Proof: The first inequality follows from Proposition 2.1.2(a)(b), since $\mathbb{R}$ is (isometric to) a subspace of B. The second follows similarly if B is finite dimensional, since every separable Banach space can be isometrically embedded in ℓ^{∞} (as is well known); the general case follows by Proposition 2.1.2(c).

Finally, by the second inequality and Proposition 2.1.3, we have that for any two σ-fields F and $G \subset M$,

$$R_{p,q}(F,G;\ell^{\infty}) = R_{q,p}(G,F;\ell^1) \leq R_{q,p}(G,F;\ell^{\infty}) = R_{p,q}(F,G;\ell^1) \leq R_{p,q}(F,G;\ell^{\infty}) \quad .$$

This implies the final equality in Theorem 2.1.5. This completes the proof.

For the next theorem, recall eqn. (0.4).

Theorem 2.1.6: For every non-trivial Banach space B the following three statements hold:

(a) $R_{1,q}(B) = R_{1,q}(\mathbb{R}) \quad \forall q,\ 1 \leq q \leq \infty$.

(b) $R_{p,1}(B) = R_{p,1}(\mathbb{R}) \quad \forall p,\ 1 \leq p \leq \infty$.

(c) In particular, $R_{1,\infty}(B) = R_{1,\infty}(\mathbb{R}) = 2\phi$.

Proof: Let us take (a) for granted for a moment, and look at (b) and (c). From (a) and Proposition 2.1.3 (and the trivial fact that $\mathbb{R}^*$ is isometric to $\mathbb{R}$) one can easily derive (b). The first equality in (c) is a special case of (a). The second equality in (c) is well known: The inequality $R_{1,\infty}(F,G;\mathbb{R}) \leq 2\phi(F,G)$ follows e.g. from [10, pp. 11-12, Lemma 1.1.9], and the opposite inequality follows from a simple argument based on the identity

$$E[I_A(I_B - I_{B^c})] - EI_A E(I_B - I_{B^c}) = 2\cdot[P(B|A)-P(B)]\cdot\|I_A\|_1 \cdot \|I_B - I_{B^c}\|_\infty$$

(say with $A \in F$, $B \in G$, $P(A) > 0$) . Now all that we need to do is prove (a).

It suffices to prove that $R_{1,q}(F,G;B) = R_{1,q}(F,G;\mathbb{R})$ in the case where $1 \leq q \leq \infty$ is arbitrary but fixed and F and G are <u>finite</u> σ-fields, say with F generated by the partition $\{A_1,\dots,A_I\}$ of Ω, and G generated by the partition $\{B_1,\dots,B_J\}$ of Ω . Without loss of generality we assume that $P(B_j) > 0$ $\forall j$.

Suppose $X \in S(F,B)$ and $Y \in S(G,B^*)$. It suffices to prove that

$$(2.1.4) \qquad |E\langle X,Y\rangle - \langle EX,EY\rangle| \leq R_{1,q}(F,G;\mathbb{R})\cdot\|X\|_1 \cdot \|Y\|_q \ .$$

Represent X and Y by $X := \Sigma_{i=1}^I x_i I_{A(i)}$ and $Y := \Sigma_{j=1}^J y_j I_{B(j)}$, with $x_i \in B$ and $y_j \in B^*$, with the events $A_1,\dots,A_I,B_1,\dots,B_J$ being as above. Then

$$|E\langle X,Y\rangle - \langle EX,EY\rangle| =$$

$$= |\Sigma_{i=1}^I \Sigma_{j=1}^J \langle x_i,y_j\rangle [P(A_i \cap B_j) - P(A_i)P(B_j)]|$$

$$\leq \Sigma_{i=1}^I \Sigma_{j=1}^J \|x_i\|_B \cdot \|y_j\|_{B^*} \cdot |P(A_i|B_j) - P(A_i)| \cdot P(B_j)$$

$$\leq [\Sigma_{j=1}^J \|y_j\|_{B^*}^q \cdot P(B_j)]^{1/q}$$

$$\cdot [\Sigma_{j=1}^J (\Sigma_{i=1}^I \|x_i\|_B \cdot |P(A_i|B_j) - P(A_i)|)^{q'} P(B_j)]^{1/q'}$$

$$= \|Y\|_q \cdot \|\Sigma_{i=1}^{I} \|x_i\|_B \cdot |E(I_{A(i)}|G) - EI_{A(i)}| \|_{q'}$$

$$\leq \|Y\|_q \cdot \Sigma_{i=1}^{I} \|x_i\|_B \cdot \|E(I_{A(i)}|G) - EI_{A(i)}\|_{q'}$$

$$\leq \|Y\|_q \cdot \Sigma_{i=1}^{I} \|x_i\|_B \cdot R_{1,q}(F,G;\mathbb{R}) \cdot \|I_{A(i)}\|_1$$

$$= R_{1,q}(F,G;\mathbb{R}) \cdot \|X\|_1 \cdot \|Y\|_q$$

by Hölder's inequality, Minkowski's inequality, and Proposition 0.1(ii). Thus eqn. (2.1.4) holds. This completes the proof.

Remark 2.1.7: We finish Section 2.1 with the following three comments:

(a) From Theorem 2.1.5, for given p and q with $1 \leq p,q \leq \infty$, one can interpret $R_{p,q}(\ell^{\infty})$ ($= R_{p,q}(\ell^1)$) as $R_{p,q}$ ("Banach space"), i.e. $\sup_B R_{p,q}(B)$.

(b) For given p and q (with $1 \leq p,q \leq \infty$) and Banach space B, one has an equality analogous to Proposition 0.1, namely

$$R_{p,q}(F,G;B) = \sup \frac{\|E(X|G) - EX\|_{q'}}{\|X\|_p}, \quad X \in S(F,B)$$

(For $X \in S(F,B)$, the definition of the q'-norm of the not-necessarily-simple B-valued function $E(X|G) - EX$ will be straightforward.)

(c) Interpolation theory applies nicely to B-valued functions; see [1, Chapter 5]. In [3, Remark 4.1] it was noted that by the Riesz (or Riesz-Thorin) interpolation theorem, there are only four equivalence classes for the measures of dependence $R_{p,q}$ in the lower triangle $1/p + 1/q \leq 1$, represented respectively by, say, $R_{\infty,\infty}$, $R_{1,\infty}$, $R_{\infty,1}$, and $R_{2,2}$. Similarly, by corresponding interpolation theorems for B-valued functions, for any given Banach space B there are at most four equivalence classes for the measures of dependence $R_{p,q}(B)$ in the lower triangle $1/p + 1/q \leq 1$, represented by, say, $R_{\infty,\infty}(B)$, $R_{1,\infty}(B)$, $R_{\infty,1}(B)$, and $R_{2,2}(B)$. By Theorem 2.1.6, $R_{1,\infty}(B)$ and $R_{\infty,1}(B)$ do not depend on the choice of (non-trivial) B; but as we shall see in the next section, $R_{\infty,\infty}(B)$ can significantly vary with B.

SECTION 2.2: ABSOLUTE REGULARITY

The measure of dependence β in eqn. (0.7) is the basis for the well known "absolute regularity" [17] condition for stochastic processes. It is already known that β is not equivalent to any of the other measures of dependence given in eqns. (0.3)-(0.6) and (0.8), or to any of the measures of dependence $\alpha_{r,s}$ ($0 \leq r,s \leq 1$) or $R_{p,q}$ ($1 \leq p,q \leq \infty$) in eqns. (0.10)-(0.11) at all. (More on this shortly.) Also, it seems that many limit theorems for real-valued (or even H-valued) r.v.'s under the "strong mixing" condition (which is based on the measure of dependence α in eqn. (0.3)) cannot be carried over to more general B-valued r.v.'s unless the assumption of strong mixing is replaced by the (stronger) assumption of absolute regularity. See e.g. the discussion and results in Dehling and Philipp [5]. To some extent, absolute regularity appears to be the most natural mixing condition for limit theory for B-valued r.v.'s. Indeed, the measure of dependence β has a quite "natural" interpretation in terms of the measures of dependence $R_{p,q}(B)$; in Theorem 2.2.1 below, it will be shown with a totally elementary argument that $\beta = (1/2)\cdot R_{\infty,\infty}(\ell^{\infty})$ ($= (1/2)\cdot R_{\infty,\infty}(\ell^{1})$) .

Now let us clarify the comparisons between β and the measures of dependence $\alpha_{r,s}$, $R_{p,q}$ in eqns. (0.10)-(0.11). Recall the inequalities $\alpha \leq \beta \leq \phi$ (see eqns. (0.3)-(0.4)). Also, β does not dominate ρ (see eqn. (0.6)). (For a σ-field $F = \{\Omega, A, A^c, \phi\}$ with $P(A) = \varepsilon$ (say with ε small), one has $\beta(F,F) = 2\varepsilon(1-\varepsilon)$ and $\rho(F,F) = 1$.) From these facts and the remarks in Section 1.2, one has that β dominates $\alpha_{r,s}$ only for $r + s < 1$, that β dominates $R_{p,q}$ only for $1/p + 1/q < 1$, and that β is dominated by $\alpha_{r,s}$ if either $r = 1$ or $s = 1$, and hence by $R_{p,q}$ if either $p = 1$ or $q = 1$. In Theorem 2.2.3 below we shall show that β is not dominated by $\alpha_{r,s}$ if $r < 1$ and $s < 1$, or even by $R_{p,q}$ if $p > 1$ and $q > 1$.

Theorem 2.2.1: $R_{\infty,\infty}(\ell^{\infty}) = R_{\infty,\infty}(\ell^{1}) = 2\beta$.

Proof: By Theorem 2.1.5 it suffices to prove that $R_{\infty,\infty}(\ell^1) = 2\beta$. Since these measures of dependence automatically satisfy eqn. (0.1), it suffices to show that if F and G are finite σ-fields then $R_{\infty,\infty}(F,G;\ell^1) = 2 \cdot \beta(F,G)$.

Suppose F is a finite σ-field, generated by a partition $\{A_1,\dots,A_I\}$ of Ω, and G is a finite σ-field, generated by a partition $\{B_1,\dots,B_J\}$ of Ω. Without loss of generality we assume that $P(A_i) > 0 \ \forall i$ and $P(B_j) > 0 \ \forall j$.

Recall that $(\ell^1)^*$ is isometric to ℓ^{∞}; we shall use the natural identification of these two spaces described in eqn. (2.1.3).

If $X \in S(F,\ell^1)$ and $Y \in S(G,\ell^{\infty})$ then one can represent these r.v.'s by $X := \Sigma_{i=1}^{I} x_i I_{A(i)}$ and $Y := \Sigma_{j=1}^{J} y_j I_{B(j)}$ (with each $x_i \in \ell^1$ and each $y_j \in \ell^{\infty}$), and one has

$$
\begin{aligned}
&|E\langle X,Y\rangle - \langle EX,EY\rangle| = \\
&= |\Sigma_{i=1}^{I} \Sigma_{j=1}^{J} \langle x_i,y_j\rangle (P(A_i \cap B_j) - P(A_i)P(B_j))| \\
&\le (\sup_i \|x_i\|_{\ell^1}) \cdot (\sup_j \|y_j\|_{\ell^{\infty}}) \cdot \sum_{i=1}^{I} \sum_{j=1}^{J} |P(A_i \cap B_j) - P(A_i)P(B_j)| \\
&= \|X\|_{\infty} \cdot \|Y\|_{\infty} \cdot 2\beta(F,G) .
\end{aligned}
\tag{2.2.1}
$$

Hence $R_{\infty,\infty}(F,G;\ell^1) \leq 2\beta(F,G)$. Conversely, in eqn. (2.2.1) equality is achieved if we take

$x_i := (0,\ldots,0,1,0,0,\ldots)$ (where the 1 is the i^{th} coordinate),

$y_j := (y_{j1},y_{j2},y_{j3},\ldots)$ where

$$y_{jk} := \begin{cases} \text{sign of } [P(A_k \cap B_j) - P(A_k)P(B_j)] & \text{if } 1 \leq k \leq I \\ 0 & \text{if } k > I \end{cases}$$

Hence $R_{\infty,\infty}(F,G;\ell^1) = 2\beta(F,G)$. This completes the proof.

Corollary 2.2.2: For every Banach space B, $R_{\infty,\infty}(B) \leq 2\beta$.

This just follows from Theorems 2.1.5 and 2.2.1 .

Theorem 2.2.3: Suppose that $1 < p,q \leq \infty$. Then (on some probability space) the following two statements hold:

(i) $R_{p,q}$ fails to dominate β .

(ii) For each p_1 and q_1 such that $1 \leq p_1,q_1 \leq \infty$, $R_{p,q}$ fails to dominate $R_{p_1,q_1}(\ell^\infty)$.

Here (ii) is a trivial corollary if (i) and Theorem 2.2.1. To prove (i) we shall show in Example 2.2.4 below that for different choices of pairs of σ-fields F and G, on different choices of probability spaces, one can (simultaneously) have $\beta(F,G) = 1/2$ and $R_{p,q}(F,G;\mathbb{R})$ arbitrarily small. Then by eqn. (0.13) and Remark 0.4, (i) will follow.

Example 2.2.4: Let p and q be arbitrary but fixed, such that $1 < p,q \leq \infty$. We shall show that $R_{p,q}(F,G;\mathbb{R})$ can be arbitrarily small for σ-fields F and G satisfying $\beta(F,G) = 1/2$. Now $R_{p,q}$ is non-increasing as p and/or q increases; without loss of generality we impose the following further restrictions on p and q:

(2.2.2) $\quad 1 < p \le 2$

(2.2.3) $\quad 1 < q < \infty$

Thus q' satisfies $1 < q' < \infty$

The following elementary fact will be needed later on:

(2.2.4) $\quad \forall\, a_1,\ldots,a_m \in \mathbb{R}, \quad (\Sigma_{k=1}^m |a_k|^p)^{1/p} \ge (\Sigma_{k=1}^m |a_k|^2)^{1/2}$.

(Let us quickly review this. Recall the elementary inequality $(a+b)^t \ge a^t + b^t$ for $a \ge 0$, $b \ge 0$, $t \ge 1$. By induction, $(\Sigma_{k=1}^m c_k)^t \ge \Sigma_{k=1}^m c_k^t$ for $t \ge 1$ and non-negative $c_1,\ldots,c_m$. Now plug in $c_k := |a_k|^p$ and $t := 2/p$, using eqn. (2.2.2).)

Now let n be an arbitrary but fixed positive integer. Construct the probability space (Ω, M, P) as follows:

Define $\Omega := \{1,2,\ldots,2n\} \times [0,1]$. Define $M =$ the family of Borel subsets of Ω. Define the probability measure P_1 on $\{1,2,\ldots,2n\}$ (with the standard discrete σ-field) by $P_1(\{j\}) = (2n)^{-1}$, $j = 1,2,\ldots,2n$. Define the probability measure P_2 on [0,1] (with the Borel σ-field) to be Lebesgue measure. Finally, define the probability measure P on (Ω, M) such that P is absolutely continuous with respect to $P_1 \times P_2$, such that

(2.2.5) $$\frac{dP}{d(P_1 \times P_2)}(j,x) := \begin{cases} 1 + h_j(x) & \text{if } 1 \le j \le n \\ 1 - h_{j-n}(x) & \text{if } n+1 \le j \le 2n \end{cases}$$

where h_1, h_2, ... are the Rademacher functions on [0,1].

Note that the marginals of P on $\{1,2,\ldots,2n\}$ and [0,1] are P_1 and P_2 respectively.

Let us define the "marginal" σ-fields

$F := \{A \times [0,1]:$ A is a subset of $\{1,2,\ldots,2n\}\}$

$G := \{\{1,2,\ldots,2n\} \times B:$ B is a Borel subset of $[0,1]\}$

By Proposition 0.1,

$$(2.2.6) \qquad R_{p,q}(F,G;\mathbb{R}) = \sup \frac{\|E(f|G) - Ef\|_{q'}}{\|f\|_p} \quad , \qquad f \in S(F,\mathbb{R})$$

Let f be an arbitrary but fixed (simple) real-valued, F-measurable r.v. We shall first obtain an upper bound on $\|E(f|G) - Ef\|_{q'} / \|f\|_p$.

This function $f = f(j,x)$ depends only on j; we can use the notation $f(j) := f(j,x) \quad \forall j = 1,2,\ldots,2n \ (x \in [0,1])$.

Define the functions f_1 and f_2 on $\{1,2,\ldots,2n\}$ as follows:

$$(2.2.7) \qquad f_1(j) := -f_1(j+n) := (f(j) - f(j+n))/2 \quad \forall j = 1,2,\ldots,n$$

$$(2.2.8) \qquad f_2(j) := f_2(j+n) := (f(j) + f(j+n))/2 \quad \forall j = 1,2,\ldots,n$$

Then $\forall j = 1,2,\ldots,2n$, $f(j) = f_1(j) + f_2(j)$.

For a.e. $x \in [0,1]$ (and every $\ell \in \{1,2,\ldots,2n\}$), we have by eqn. (2.2.5) and some simple calculations, $E(f|G)(\ell,x) - Ef = \Sigma_{j=1}^n f_1(j) \cdot h_j(x)/n$. Therefore

$$\|E(f|G) - Ef\|q' = (1/n) \cdot [\int_0^1 | \Sigma_{j=1}^n f_1(j) \ h_j(x)|^{q'} dx]^{1/q'}$$

By eqn. (2.2.7) and the convexity of the function u^p (for nonnegative u), we have that $\|f_1\|_p \leq \|f\|_p$. Also $\|f_1\|_p = [(1/n) \cdot \Sigma_{j=1}^n |f_1(j)|^p]^{1/p}$ by a trivial calculation. Hence by eqn. (2.2.4) and Khinchin's inequality (see e.g. [20, Chapter 5, Theorem (8.4)]) ,

$$\frac{\|E(f|G) - Ef\|_{q'}}{\|f\|_p} \leq \frac{n^{(1/p)-1}\ [\int_0^1 |\Sigma_{j=1}^n\ f_1(j)\ h_j(x)|^{q'} dx]^{1/q'}}{[\ \Sigma_{j=1}^n\ |f_1(j)|^p]^{1/p}}$$

$$\leq \frac{n^{(1/p)-1}\ C(q')\ \cdot\ [\Sigma_{n=1}^n\ (f_1(j))^2]^{1/2}}{[\ \Sigma_{j=1}^n\ |\ f_1(j)\,|^2]^{1/2}}$$

$$= n^{-1/p'} \cdot C(q')$$

where the constant $C(q')$ depends only on (our fixed value of) q'.

Since f was arbitrary, we have by eqn. (2.2.6),

$$R_{p,q}(F\ G;\mathbb{R}) \leq n^{-1/p'} \cdot C(q')$$

By making n sufficiently large to begin with, we can make $R_{p,q}(F,G;\mathbb{R})$ arbitrarily small. Also, by eqn. (2.2.5) and an elementary argument, regardless of the value of n,

$$\beta(F,G) = \frac{1}{2} \cdot \int_\Omega \left|\frac{dP}{d(P_1 x P_2)} - 1\right| d(P_1 \times P_2)$$

$$= \frac{1}{2} \cdot \int_\Omega 1\ d(P_1 \times P_2)$$

$$= \frac{1}{2}$$

This completes the discussion for Example 2.2.4, and thereby completes the proof of Theorem 2.2.3.

Remark 2.2.5: Gastwirth and Rubin [7] defined for each t, $1 \leq t \leq \infty$, the following measure of dependence:

$$\Delta_t(F,G) := \|\mathrm{var}[P_G(\cdot\,|\ F) - P_G(\cdot)]\|_t$$

where P_G is the restriction of the measure P to events in G. For finite σ-fields F and G the definition of $\Delta_t(F,G)$ is clear; and for general σ-fields F and G, where a measure-theoretic ambiguity may have to be settled, we shall simply stipulate that eqn. (0.1) be satisfied by Δ_t. It is easy to see that $\Delta_1 = 2\beta$, $\Delta_\infty = 2\phi$, and that Δ_t is equivalent to β if $1 < t < \infty$. By a simple argument similar to that of Theorems 2.1.6 and 2.2.1, one has that for each t, $1 \le t \le \infty$,

$$\Delta_t = R_{t',\infty}(\ell^\infty) = R_{t',\infty}(\ell^1).$$

SECTION 2.3: THE HILBERT SPACE CASE

In [3, Section 4.3] measures of dependence based on H-valued (Hilbert space valued) random variables were examined. Here we shall extend the main result of that section. As earlier it suffices to consider real spaces.

Let H be an arbitrary real Hilbert space, with inner product $\langle\cdot,\cdot\rangle$.

For each p, $1 \le p < \infty$, the quantity $Z_p := [(2\pi)^{-\frac{1}{2}}\int_{-\infty}^{\infty}|x|^p \exp(-x^2/2)dx]^{1/p}$ (the p-norm of a standard $N(0,1)$ r.v.) will be useful.

Theorem 2.3.1. If H is a non-trivial Hilbert space, $(p,q) \in [1,\infty]^2$, and $F,G \subset M$, then $R_{p,q}(F,G;H) \le A \cdot R_{p,q}(F,G)$ where $A = A(p,q)$ is a function only of p and q. Further, if $1 \le p,q < \infty$ then one can take $A(p,q) = Z_{p_1} \cdot Z_{q_1}$ for any choice of p_1 and q_1 satisfying $p \le p_1 < \infty$, $q \le q_1 < \infty$, and $p_1^{-1} + q_1^{-1} \le 1$; and in particular, if $1 \le p,q \le 2$ then $R_{p,q}(F,G;H) = R_{p,q}(F,G;\mathbb{R})$.

Thus for a given $(p,q) \in [1,\infty]^2$ the measure of dependence $R_{p,q}(H)$ is linearly equivalent to $R_{p,q}(\mathbb{R})$ (or to $R_{p,q}$ in eqn. (0.10)). For the case $p = q = \infty$, this was shown by Dehling and Philipp [5, p. 692, Lemma 2.2], who used the Grothendieck inequality in their proof. Under the restriction $p^{-1} + q^{-1} \le 1$ one has Theorem 2.3.1 from [3, Theorem 4.2]; the purpose here is to remove that extra restriction. Theorem 2.3.1 is not a particularly new result, but rather an adaptation to our context of results well known in functional analysis.

The very last case of Theorem 2.3.1 $(1 \le p,q \le 2)$ follows from the rest of that theorem by taking $p_1 = q_1 = 2$. (See eqn. (2.3.1) below.)

The proof of Theorem 2.3.1 will be based on the following lemma:

Lemma 2.3.2. Suppose $1 \le p \le p_1 \le \infty$, $1 \le q \le q_1 \le \infty$, and $p_1^{-1} + q_1^{-1} \le 1$. Suppose $X_1, \ldots, X_I$ and $Y_1, \ldots, Y_J$ are real-valued (or complex-valued) random variables. Then

$$E(\|(X_i)_{i=1}^I\|_{\ell^p} \cdot \|(Y_j)_{j=1}^J\|_{\ell^q}) \le \|(\|X_i\|_{p_1})_{i=1}^I\|_{\ell^p} \cdot \|(\|Y_j\|_{q_1})_{j=1}^J\|_{\ell^q}$$

Proof.

$$E(\|(X_i)_{i=1}^I\|_{\ell^p} \cdot \|(Y_j)_{j=1}^J\|_{\ell^q}) \le \|\,\|(X_i)_{i=1}^I\|_{\ell^p}\|_{p_1} \cdot \|\,\|(Y_j)_{j=1}^J\|_{\ell^q}\|_{q_1}$$

$$\le \|(\|X_i\|_{p_1})_{i=1}^I\|_{\ell^p} \cdot \|(\|Y_j\|_{q_1})_{j=1}^J\|_{\ell^q}$$

Here the first inequality comes from Hölder's inequality (and the fact that our measure is a probability measure). To show the second inequality it is enough to show that

$$\|\,\|(X_i)_{i=1}^I\|_{\ell^p}\|_{p_1} \le \|(\|X_i\|_{p_1})_{i=1}^I\|_{\ell^p} .$$

If $p_1 = \infty$ then this is simple. (Also, if $p = 1$ then this is simply Minkowski's inequality.) If $p_1 < \infty$ then

$$\|\,\|(X_i)_{i=1}^I\|_{\ell^p}\|_{p_1} = [E(\sum_{i=1}^I |X_i|^p)^{p_1/p}]^{1/p_1}$$

$$= [\|\sum_{i=1}^I |X_i|^p\|_{p_1/p}^{p_1/p}]^{1/p_1} = [\|\sum_{i=1}^I |X_i|^p\|_{p_1/p}]^{1/p}$$

$$\leq \left[\sum_{i=1}^{I} \| \, |X_i|^p \|_{p_1/p} \right]^{1/p} = \left[\sum_{i=1}^{I} [E|X_i|^{p \cdot p_1/p}]^{p/p_1} \right]^{1/p}$$

$$= \left[\sum_{i=1}^{I} \|X_i\|_{p_1}^{p} \right]^{1/p} = \| (\|X_i\|_{p_1})_{i=1}^{I} \|_{\ell^p} .$$

This completes the proof of Lemma 2.3.2.

Proof of Theorem 2.3.1. For the cases $p = \infty$ or $q = \infty$, see [3, Theorem 4.2]. Here we only consider the case where $1 \leq p < \infty$ and $1 \leq q < \infty$. As in the proof of [3, Theorem 4.2] we restrict our attention to an arbitrary finite-dimensional real Hilbert space H; we use the same Gaussian probability measure γ on H as was used there. (The use of a Gaussian measure here is similar to its use in Rietz [14].)

Suppose $X := \Sigma_{i=1}^{I} x_i I_{A(i)}$ and $Y := \Sigma_{j=1}^{J} y_j I_{B(j)}$ where $\{A_1, \dots, A_I\}$ and $\{B_1, \dots, B_J\}$ are each a partition of Ω with $A_i \in \mathcal{F}$ for all i and $B_j \in \mathcal{G}$ for all j, and $x_1, \dots, x_I$ and $y_1, \dots, y_J$ are elements of H. To prove Theorem 2.3.1 it suffices to prove that

$$(2.3.1) \qquad |E\langle X,Y\rangle - \langle EX,EY\rangle| \leq R_{p,q}(\mathcal{F},\mathcal{G};\mathbb{R}) \cdot Z_{p_1} \cdot Z_{q_1} \cdot \|X\|_p \cdot \|Y\|_q$$

holds for every choice of p_1 and q_1 meeting the specifications $p \leq p_1 < \infty$, $q \leq q_1 < \infty$, and $p_1^{-1} + q_1^{-1} \leq 1$.

Before we prove eqn. (2.3.1), a couple of preliminary observations will be needed. For any $x, y \in H$ one has $\langle x,y\rangle = \int_H \langle x,u\rangle \cdot \langle y,u\rangle \gamma(du)$. (This was used in the proof of [3, Theorem 4.2].) For any $x \in H$ the r.v. $V: H \to \mathbb{R}$ defined by $V(u) := \langle x,u\rangle$ is a $N(0, \|x\|_H^2)$ r.v. (on the probability space (H,γ)), and hence for any t, $1 \leq t < \infty$, $\int_H |\langle x,u\rangle|^t \gamma(du) = \|x\|_H^t \cdot Z_t^t$. These observations, Fubini's theorem, and Lemma 2.3.2 (on the probability space (H,γ)) will now be used to prove eqn. (2.3.1), under the given specifications on p_1 and q_1, as follows:

$$
\begin{aligned}
&|E\langle X,Y\rangle - \langle EX,EY\rangle| \\
&\quad = \Big|\int_H E(\langle X,u\rangle\cdot\langle Y,u\rangle)\gamma(du) - \int_H \langle EX,u\rangle\cdot\langle EY,u\rangle\gamma(du)\Big| \\
&\quad = \Big|\int_H [E(\langle X,u\rangle\cdot\langle Y,u\rangle) - E\langle X,u\rangle\cdot E\langle Y,u\rangle]\gamma(du)\Big| \\
&\quad \le \int_H R_{p,q}(F,G;\mathbb{R})\cdot[E|\langle X,u\rangle|^p]^{1/p}\cdot[E|\langle Y,u\rangle|^q]^{1/q}\gamma(du) \\
&\quad = R_{p,q}(F,G;\mathbb{R})\int_H \|(\langle x_i,u\rangle\cdot P^{1/p}(A_i))_{i=1}^I\|_{\ell^p}\cdot\|(\langle y_j,u\rangle\cdot P^{1/q}(B_j))_{j=1}^J\|_{\ell^q}\gamma(du) \\
&\quad \le R_{p,q}(F,G;\mathbb{R})\cdot\|((\int_H|\langle x_i,u\rangle\cdot P^{1/p}(A_i)|^{p_1}\gamma(du))^{1/p_1})_{i=1}^I\|_{\ell^p} \\
&\qquad\qquad \cdot\|((\int_H|\langle y_j,u\rangle\cdot P^{1/q}(B_j)|^{q_1}\gamma(du))^{1/q_1})_{j=1}^J\|_{\ell^q} \\
&\quad = R_{p,q}(F,G;\mathbb{R})\cdot\|(P^{1/p}(A_i)\cdot(\int_H|\langle x_i,u\rangle|^{p_1}\gamma(du))^{1/p_1})_{i=1}^I\|_{\ell^p} \\
&\qquad\qquad \cdot\|(P^{1/q}(B_j)\cdot(\int_H|\langle y_j,u\rangle|^{q_1}\gamma(du))^{1/q_1})_{j=1}^J\|_{\ell^q} \\
&\quad = R_{p,q}(F,G;\mathbb{R})\cdot\Big[\sum_{i=1}^I P(A_i)\cdot(\|x_i\|_H^{p_1}\cdot Z_{p_1}^{p_1})^{p/p_1}\Big]^{1/p} \\
&\qquad\qquad \cdot\Big[\sum_{j=1}^J P(B_j)\cdot(\|y_j\|_H^{q_1}\cdot Z_{q_1}^{q_1})^{q/q_1}\Big]^{1/q} \\
&\quad = R_{p,q}(F,G;\mathbb{R})\cdot Z_{p_1}\cdot Z_{q_1}\cdot\|X\|_p\cdot\|Y\|_q
\end{aligned}
$$

Thus eqn. (2.3.1) holds. This completes the proof of Theorem 2.3.1.

Acknowledgments. The authors thank A. Gillespie for calling their attention to Zafran [19], and A. Gut for his interest and encouragement.

REFERENCES

1. J. Bergh and J. Löfström. Interpolation Spaces. Springer-Verlag, Berlin, 1976.

2. R.C. Bradley. Equivalent measures of dependence. J. Multivar. Anal. 13 (1983) 167-176.

3. R.C. Bradley and W. Bryc. Multilinear forms and measures of dependence between random variables. J. Multivar. Anal. (to appear).

4. A.V. Bulinskiĭ. On measures of dependence close to the maximal correlation coefficient. Soviet Math. Dokl. 30 (1984) 249-252.

5. H. Dehling and W. Philipp. Almost sure invariance principles for weakly dependent vector-valued random variables. Ann. Probab. 10 (1982) 689-701.

6. M. Denker and G. Keller. On U-statistics and v. Mises' statistics for weakly dependent processes. Z. Wahrsch. verw. Gebiete 64 (1983) 502-522.

7. J.L. Gastwirth and H. Rubin. The asymptotic distribution theory of the empiric c.d.f. for mixing stochastic processes. Ann. Statist. 3 (1975) 809-824.

8. R.A. Hunt. On L(p,q) spaces. Enseignement Math. 12 (1966) 249-276.

9. I.A. Ibragimov and Y.V. Linnik. Independent and Stationary Sequences of Random Variables. Wolters-Noordhoff, Groningen, 1971.

10. M. Iosifescu and R. Theodorescu. Random Processes and Learning. Springer-Verlag, Berlin, 1969.

11. B.A. Lifshits. Invariance principle for weakly dependent variables. Theory Probab. Appl. 29 (1984) 33-40.

12. A. Marshal and I. Olkin. Inequalities: Theory of Majorization and its Applications. Academic Press, New York, 1979.

13. M. Peligrad. A note on two measures of dependence and mixing sequences. Adv. Appl. Probab. 15 (1983) 461-464.

14. R.E. Rietz, A proof of the Grothendieck inequality. Israel J. Math 19 (1974) 271-276.

15. M. Rosenblatt. Markov Processes, Structure and Asymptotic Behavior. Springer-Verlag, Berlin, 1971.

16. E.M. Stein and G. Weiss. An extension of the theorem of Marcinkiewicz and some of its applications. J. Math. Mech. 8 (1959) 263-284.

17. V.A. Volkonskii and Y.A. Rozanov. Some limit theorems for random functions I. Theory Probab. Appl. 4 (1959) 178-197.

18. C.L. Withers. Central limit theorems for dependent random variables I. Z. Wahrsch. verw. Gebiete 57 (1981) 509-534.

19. M. Zafran. A multilinear interpolation theorem. Studia Math. 62 (1978) 107-124.

20. A. Zygmund. Trigonometric Series, Volumes I and II. Cambridge University Press, Cambridge, 1959.

ADDENDUM.

In connection with Theorem 1.1.1 and [3, Theorem 3.6], here are some extra comments, added when this report was in the final stages of its preparation. The multilinear form B in Theorem 1.1.1 automatically has the property that $\frac{1}{2}B$ is a "product form" (in the terminology used in [3, Theorem 3.6]). For any given $\underline{p}$, for $t > 0$ sufficiently small one can refine Theorem 1.1.1 by having $d_{\underline{p}}(B) = t$, with B itself being a product form, (without disturbing the other properties in Theorem 1.1.1). (The positive constant $C = C(\underline{p})$ may have to be made smaller, but that is of no significance.) To carry out this refinement, we can restrict our attention to the case where $1 < p_k < \infty$ for some k, as in the proof of Theorem 1.1.1. By examining the bottom four lines of p. 12 (and the two top lines of p. 13) in the case where $B_k = [0,\frac{1}{2}]$ for all k, we see that $d_{\underline{p}}(B) \geq 2^{-n+1}t$ in that construction. For $t > 0$ sufficiently small (depending on $\underline{p}$) we can start by simply using that construction with t replaced by $2^n t$. Then $2t \leq d_{\underline{p}}(B) \leq 2^n t$. To then modify the construction so that $d_{\underline{p}}(B) = t$ and B is a product form (without disturbing the other properties in Theorem 1.1.1) we simply define the number $a := t/d_{\underline{p}}(B)$ and then replace the probability measure P in the construction by the new probability measure $a \cdot P + (1 - a) \cdot$[Lebesgue measure on $[0,1]^n$].

This refinement can also be easily worked out with B being the n-dimensional cumulant, because of the fact (see p. 12, line 7) that the (n - 1)-dimensional marginals of the given probability measure P are simply (n - 1)-dimensional Lebesgue measure. If one isn't concerned about the particular formula for the multilinear form B, then (for a given $\underline{p}$) "not so small" values of $t \leq 1$ can also be covered in this refinement by simply using the multilinear form $B(f_1, \ldots, f_n) = t \cdot E(f_1 \cdot \ldots \cdot f_n)$ on a trivial probability space consisting of just a single point. Also, Remark 1.1.2 can be refined (with an appropriate linear functional) in an exactly analogous manner as above.

END

FILMED

11-85

DTIC